DOCUMENT RESUME

ED 181 658 EC 122 722

AUTHOR	Black, Talbot, Ed.
TITLE	Perspectives on Measurement: A Collection of Readings for Educators of Young Handicapped Children. 1979 Series, No. 1.
INSTITUTION	North Carolina Univ., Chapel Hill. Technical Assistance Development System.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
PUB DATE	79
CONTRACT	300-77-0507
NOTE	89p.; This book is a Proceedings Document based on the Measuring Child Progress Conference held by TADS in Nashville, TN
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Early Childhood Education; *Evaluation Methods; *Handicapped Children; Infants; *Measurement Techniques; Preschool Education; *Program Evaluation; Screening Tests; Severely Handicapped; Student Evaluation; Tests; *Test Selection

ABSTRACT

Eleven author contributed papers presented at a 1978 conference on measuring the growth and development of young handicapped children are presented. N. Anastasiow presents a rationale for measurement in his paper "A Philosophical Perspective: Why Measure Child Progress?'" Seven papers address instrument selection: "Screening, Diagnosis, and Assessment: How Do These Types of Measurement Differ?" (M. Kemper and W. Frankenburg); "Mildly to Moderately Handicapped Preschoolers: How Do You Select Child Assessment Instruments?" (G. Harbin); "Mildly to Moderately Handicapped Preschoolers: How Do You Select Program Evaluation Instruments?" (R. Sheehan); "Mildly to Moderately Handicapped Infants: What Should Influence Your Approach To Measurement?" (C. Kopp); "Mildly to Moderately Handicapped Infants: How Do You Assess the Progress of Both Mother and Infant?" (E. Badger); "Severely and Profoundly Handicapped: How Do You Measure Sensorimotor and Cognitive Skills?" (C. Robinson); and "Severely and Profoundly Handicapped: How Do You Measure Communication Skills?" (R. Dubose). Three final papers conclude the volume: "Public Law 94-142: What Are Its Implications for Measurement?" (G. Harbin); "Adaptations in Measurement Procedures: Should You Make Alterations for Handicapped Children?" (R. DuBose); and "Documenting Child Progress: How Do You Choose Data Collection Designs and Instruments?" (R. Sheehan). (CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED181658

PERSPECTIVES ON MEASUREMENT

A COLLECTION OF READINGS FOR EDUCATORS OF YOUNG HANDICAPPED CHILDREN

EDITED BY TALBOT BLACK

1979 SERIES NUMBER 1

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Faith J. Meyer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED 182122

Contributors

Nicholas Anastasiow, formerly Director of the Institute for Child Study, Indiana University, Bloomington, Indiana. Presently, Director of Research, John F. Kennedy Center for Child Development, University of Colorado Health Sciences Center, Denver, Colorado.

Earladeen Badger, Director of Infant Stimulation/Mother Training Project, University of Cincinnati College of Medicine, Department of Pediatrics, Cincinnati, Ohio.

Rebecca Fewell DuBose, Associate Professor, Experimental Education Unit, University of Washington, Seattle, Washington.

William K. Frankenburg, Director, John F. Kennedy Center for Child Development, University of Colorado Health Sciences Center, Denver, Colorado.

Gloria Harbin, formerly Early Childhood Educational Program Specialist for the Technical Assistance Development System (TADS) and the Mid-East Learning Resources System (MELRS). Presently completing her doctoral degree at the University of North Carolina, Chapel Hill, North Carolina.

Melinda B. Kemper, Staff Psychologist, John F. Kennedy Center for Child Development, University of Colorado Health Sciences Center, Denver, Colorado.

Claire B. Kopp, Associate Professor in Residence at the University of California at Los Angeles and Co-Director of Project REACH, Los Angeles, California.

Cordelia C. Robinson, Director, Infant Development Program, University of Nebraska Medical Center, Omaha, Nebraska.

Robert Sheehan, on leave from the Department of Curriculum and Instruction, University of Virginia at Charlottesville. Presently a Post Doctoral Fellow at Project REACH, University of California Graduate School of Education, Los Angeles, California.

Perspectives on Measurement
A Collection of Readings for Educators of Young Handicapped Children

edited

by

Talbot Black

This Book is a Proceedings Document based on the Measuring Child Progress Conference held by TADS in Nashville, Tennessee.

Published by the Technical Assistance Development System (TADS), a division of the Frank Porter Graham Child Development Center, at the University of North Carolina at Chapel Hill

Summer 1979

SPECIAL THANKS

A special thank you is extended to–Tanya Suarez, Pat Vandtviere, Sonya Prestridge, Jim Cox, Mike Woodard, David Lillie, and Pascal Trohanis for their advice and editorial suggestions in the preparation of this book.

This book was developed pursuant to contract 300-77-0507 from the United States Office of Education. Contractors who undertake such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

BEH Project Officer, Dr. David Rostetter

Cover logo designed by Nancy Enlow
Text type composed by Mary N. Watkins
Managing editor, Kennith W. Goin

The text of this book is 10 point IBM Times Roman.

Contents

Preface . . . *vii*

SECTION I: RATIONALE

Chapter 1 *A Philosophical Perspective:* . . . 3
Why measure child progress?
Nicholas Anastasiow

SECTION 2: INSTRUMENT SELECTION

Chapter 2 *Screening, Diagnosis, and Assessment:* . . . 11
How do these types of measurement differ?
Melinda B. Kemper and William K. Frankenburg

Chapter 3 *Mildly to Moderately Handicapped Preschoolers:* . . . 20
How do you select child assessment instruments?
Gloria Harbin

Chapter 4 *Mildly to Moderately Handicapped Preschoolers:* . . . 29
How do you select program evaluation instruments?
Robert Sheehan

Chapter 5 *Mildly to Moderately Handicapped Infants:* . . . 32
What should influence your approach to measurement?
Claire B. Kopp

Chapter 6 *Mildly to Moderately Handicapped Infants:* . . . 39
How do you assess the progress of both mother and infant?
Earladeen Badger

Chapter 7 *Severely and Profoundly Handicapped:* . . . 47
How do you measure sensorimotor and cognitive skills?
Cordelia C. Robinson

Chapter 8 *Severely and Profoundly Handicapped:* . . . 53
How do you measure communication skills?
Rebecca Fewell Dubose

SECTION 3: OTHER ISSUES

Chapter 9 *Public Law 94-142:* . . . 61
What are its implications for measurement?
Gloria Harbin

Chapter 10 *Adaptations in Measurement Procedures:* . . . 75
Should you make alterations for handicapped children?
Rebecca Fewell DuBose

Chapter 11 *Documenting Child Progress:* . . . 79
How do you choose data collection designs and instruments?
Robert Sheehan

Preface

This publication is the product of a topical conference prepared and conducted by the Technical Assistance Development System (TADS) for the Handicapped Children's Early Education Program (HCEEP).

The conference, *Measuring Progress in Young Handicapped Children*, was held in Nashville, Tennessee, on September 12-14, 1978. It was attended by eighty-three participants representing thirty demonstration projects and twenty-two outreach projects. The topic was selected because of its growing importance, both within the HCEEP network and the general field of early childhood special education.

Rationale for this Book

The need to measure the growth and development of young handicapped children in order to demonstrate the effectiveness of early intervention has always been a primary concern of HCEEP. However, recent developments have provided new incentives for exploring the issues involved in measurement:

1) **The field of early childhood special education is coming of age.** After ten years of HCEEP, there are large numbers of professionals in special education and other related fields for whom early education of the handicapped has become a long-term career focus. Today, a broad array of exemplary programs in early education of the handicapped are operating across the country, including demonstration and outreach projects, training programs, and major research efforts. Many states are now requiring services for young handicapped children through laws, and still others have initiated early childhood programs without legislative mandates. Committees and consortia are being formed by professionals in the field at state, regional, and national levels in order to share best practices, discuss issues of common concern, and speak in a collective voice for their special interests. Within HCEEP, there is a national committee concerned with infant programs, a consortium of programs serving rural areas, a committee concerned about minority children and families, and several state consortia. Early childhood is becoming more prominent on the agenda of CEC conventions . . . and we could go on. The point is that there is now a critical mass of people and programs that are ready and capable of refining and improving our current methods of measuring the progress of young handicapped children.

2) **There is a call for accountability.** With the increased visibility of early childhood special education has come increased expectations. We encounter more questions like, "How do you know the children in your program are improving?" or "How do you know for sure it was your program's intervention that helped them improve?" or even "How do you know that the children in your program are really handicapped?" We are hearing these questions more frequently, and not just from our colleagues in special education and the behavioral sciences. We also hear them from school boards, legislators, and parents. Those who now hold us accountable—including our peers—expect to see clear evidence provided by proven and dependable measurement procedures.

   Some of these expectations are not entirely realistic. They often reflect a lack of appreciation for the unique measurement problems that come with young handicapped children. Nevertheless, the call for accountability must be answered by early childhood special educators if we are going to maintain the momentum our field has enjoyed over the last ten years, to continue competing successfully for scarce resources, and to expand services to include all handicapped children.

3) **New Legislation for the handicapped has raised measurement issues.** Public Law 94-142 and other state and federal laws concerned with the handicapped are finally beginning to provide long-needed rights and benefits, but they also require that judgments and decisions be made by those responsible for implementing the laws. Many of those judgments involve the use of testing and measurement procedures: such as casefinding and screening, diagnostic evaluations in order to determine placement, and IEP development and review. When the handicapped person in question happens to be a young child, those involved will look to early childhood special education for the

instruments and procedures upon which to base their decisions. The wisdom of their decisions and the consequences for the child will be determined largely by the quality and usefulness of what we in the field can provide.

Purposes and Audiences for this Book

The HCEEP Network is perhaps in the best position of any group to provide leadership for improving measurement procedures for young handicapped children because of the ten years of experience and the human resources it contains. One purpose of this publication is to emphasize the importance of this topic for HCEEP. The second and more immediate purpose is to provide technical assistance to HCEEP project staffs, especially the project director, in the area of child measurement. Project staff members must make several decisions about how the children in their project are measured for different purposes, including: screening, placement, individualized education programming, and, of course, how children's progress will be measured to demonstrate program effectiveness.

A project staff must consider many issues in deciding what instruments and procedures will be used (or developed) by the project such as, "Can we adapt a standardized test and still use the results in our evaluation?" or "How does PL 94-142 affect the way we test children?" The chapters in this publication identify and discuss many such questions. We hope our readers will gain a broader understanding of the issues and, as a result, be able to make more informed decisions regarding how children are measured in their programs.

About this Book

The publication is divided into three sections. *Section 1* - the conference key-note speech - offers a rationale for why we should measure child progress. *Section 2* contains a series of chapters, organized around a conceptual framework, on the ways in which children are measured in a typical early childhood program. *Section 3* discusses three issues that tend to cut across various types of programs.

Several limitations of this publication warrant mention. The reader should recognize at the outset that this is a "proceedings document"; i.e., it attempts to convey what was presented at the conference. It does not, consequently, contain a consensus on what constitutes the best measurement practice. The speakers were invited because of their experience and expertise, not because they all agreed with each other. In fact, a diversity of views was welcomed since the topic itself is in its formative stage.

This publication is by no means a full and comprehensive treatment of the topic; rather it is a limited and focused discussion of issues. For example, program evaluation is dealt with in a limited way by focusing largely on quantitative data collection strategies. The full range of alternative evaluation designs and strategies is not presented here. These limitations are due, in part, to the constraints imposed by the conference itself. One can only cover so much in a two-day conference. The common goal of all the chapters is to provide information to HCEEP project staffs for decision making.

Closing

As a result of the conference and through editing this publication, I have formulated an impression of where we, as a field, are in being able to measure the growth and development of young handicapped children. Let me share my impression through the following analogy. As educators of young handicapped children, we are like the lawyer seeking to defend his client's cause. In building his case, the lawyer longs for an eye-witness, an expert, an unimpeachable source. But alas, as yet he has found none. So for now, the lawyer must struggle with the more difficult and less certain task of building his case on circumstantial evidence. And while no single piece of evidence by itself can prove his case beyond the shadow of a doubt, the preponderance of all the evidence will hopefully carry the day for his client's cause.

In early childhood special education, we have no single test or instrument or set of procedures or evaluation design that can be our "unimpeachable source." We must struggle with what is currently available and often base our case on admittedly circumstantial evidence, whether that case be proving the effectiveness of our early intervention efforts or simply making judgments about where to place a young handicapped child and how to program for his or her needs. The point is that what is currently available can, in most respects, serve us well. We need not be overly critical and apologetic about the current "state of the art." Nevertheless, we in HCEEP continue to have, I believe, a special obligation to improve our instruments and procedures and to share those with the field. Who knows? Maybe some of us will find an expert eye-witness or two after all.

SECTION 1: RATIONALE

CHAPTER 1
A Philosophical Perspective: Why measure child progress?
by Nicholas Anastasiow

Introduction

This chapter is designed for persons who work in programs for preschool handicapped children. Its aim is to raise questions about the purposes for measuring progress in terms of evaluation and program development. The ideas which follow are very straight forward and have been around for a long time. As you read, you may even pause to think, "I already knew that." Hopefully, you will. More importantly, we hope that you will find the reasons for measuring pupil progress that we have listed to be truths which are simple, but not simplistic.

Program evaluation has become an integral part of educational endeavors. Although ubiquitous, its form and purpose vary widely from place to place. Variation per se is not undesirable, but in the case of evaluation there are so many preconceptions by both *lay public* and *decision makers* about what the end results of evaluation should be, that many persons who are involved in program development and mandated to "evaluate" their efforts, approach evaluation with an uneasiness bordering on fear and often open hostility and anger. In many instances, there are some very good reasons for the uneasiness, and equal justification for the anger. As Glass (1975) has indicated, evaluation has not always been used fairly. In the section that follows, the varying definitions of evaluation and its purposes will be discussed in attempting to delineate the sources of the confusion.

Lack of Agreement

One of the major problems in the field of evaluation (and therefore in measuring pupil progress) is the lack of agreement among the major evaluation theorists as to the central purpose of evaluation.

Stengel (1978) notes that evaluation specialists see evaluations serving government and politics (Weiss, Brickell); research (Guba); decision makers (Alkin, Sommer); curriculum improvement (Cronbach); parents (Wolfe); and the political system as value enforcer (Karier).

In general, the theorists do agree that evaluation should be conducted. As Stengel (1978) states, part of this consensus is based on a common agreement among Americans that society can be improved and that its members can learn how to be effective. However, as Stengel indicates, Americans do not agree on their definition of effectiveness. I would add that Americans also do not agree on how to go about helping others become more effective. At this broad level, the press to evaluate becomes a force to both define effectiveness and to demonstrate that the educational program is instrumental in helping others become effective. To do the first (define effectiveness) requires statements of goals and objectives. To demonstrate the latter (demonstrate that the program is effective) requires that one engage in the entire process of defining goals, objectives, procedures, behaviors, program processes, and on and on (Anastasiow, 1973).

The major point is that most decision makers treat the definition of effectiveness and the proof that an educational program is effective as two simple tasks that common sense alone would dictate you ought to do and ought to be able to do. Yet, both tasks are extremely difficult and require a great deal of thought and effort. What must be kept in mind is that a definition of effectiveness is based on one's philosophy of man and one's value system. What one measures and how one measures it must be consistent with the underlying reasons for engaging in the activity in the first place. Most of us generally assume we have agreement when, in fact, there are wide-spread differences regarding value orientations.

For example, many decision makers will insist on it being proven that your program has some utility or use before they will continue support. The utility or use may be expressed in terms of cost effectiveness. Whereas, program personnel feel that providing services to a handicapped child is of such intrinsic value that one ought not have to justify it. I believe these two value orientations are often in conflict and House (1976) recently has written very cogently about them. House differentiates between the utilitarian and libertarian orientations in the United States and comes out strongly for the right of the individual. Most of us would agree with House, but we should not be so naive as to assume that the right of the individual is the law of the land. An analysis of Supreme Court decisions clearly demonstrates the tendency to favor the rights of property over the rights of the individual (Kluger, 1975). The evaluator who takes a stance favoring individual rights must take great care to insure that his methods do not violate those rights.

A Transactional View of Learning

In what follows, a value system based on the worth of the individual in the context of a transactional view of learning is examined. Unfortunately, transactional systems

are complex and difficult to define. In the transactional approach selected here, each individual in the system is viewed as developing. Development is considered a life-span adventure with major discontinuities due to physical maturation, and later, to the wearing down of physical systems (such as vision). Within this orientation all persons are learners (child, parents, teachers, staff), i.e., by actively participating, they influence the outcomes of the learning.

"How one learns" is a more intriguing subject than "does one learn." Insights into the process of how one learns are in short supply, but the study of this process may be our best contribution to the field. As Cronbach (1975) recently stated, the goal of science is not to amass generalizations atop a theoretical tower, but to gain insight into contemporary relationships through "intensive local observations which (go) beyond any discipline into an open eyed, open minded appreciation of the surprises nature deposits into the investigator's net" (p. 12).

Most of what follows revolves around Cronbach's notion of the need for careful, intensive, and continuous observation of what is happening. That we will see only what we intend to see and give meaning only to that upon which we focus is an often unrecognized phenomenon. Laurence Darell (1957) states it nicely, "Does not everything depend upon our interpretation of the silence around us" (p. 245)?

A Developmental Perspective

House's (1976) notions of justice in evaluation highlight the need for evaluation to serve children and staff as well as decision makers. This focus on the child reflects the current trend in evaluation towards the need to provide data to all the participants, not just the funding agency and its constituents, Congress and the White House.

However, focusing on the child's needs in the measurement of pupil progress raises some very difficult questions which are not answered when group mean change data are used to measure program impact. For example, group pre-post designs which measure IQ as an indication of program impact mask changes in an individual child. These designs have been seriously criticized on a number of dimensions (see Takanishi (1978) for an excellent summary critique of the evaluation of early childhood programs). As Takanishi (1978) notes, change is too often defined in terms of test score change rather than attending to the structures underlying change [that is the child's maturation or acquisition of cognitive strategies]. The field of child development overall has been criticized for its lack of developmental perspective in research (McCall, 1977). Rarely does one find longitudinal studies which measure changes of the same children such as Werner and Smith (1977).

Several research workers and program developers are attempting to construct techniques to approach a developmental perspective [see particularly Takanishi (1978) and Magoon (1977) for overviews]. Riegel (1972) argues strongly for a multi-dimensional approach to the evaluation of a child, and Brych, Meisels and Markowitz (1979) argue for a developmental perspective in evaluating integration of handicapped children into an open framework preschool.

If we take seriously Kohlberg and Mayer's (1972) notion of "Development as the Aim of Education," then developmental processes are what need to be focused on in data collection. It should be recognized, however, that a developmental approach based on cognitive models poses some difficult measurement problems (Guralnick, 1973; Brych, Meisels and Markowitz, 1979; Guba, 1969).

Perhaps foremost is Takanishi's (1978) statement that if the programs are designed "to enhance the child's life chances," then the individual is the only unit of analysis that makes any sense.

Audiences for Pupil Progress

Why does measuring pupil progress serve so many goals in a program? Whether your focus be advising, informing, improving or responding to decision makers, each person in the program hierarchy has use of the data. The sections that follow focus on the special needs of each person in your system but do not preclude use of the same set of data for some other purpose. For example, Stake (1967) would see data serve the teacher, by helping his or her teacher refine techniques, whereas Tyler (1978) may perceive the same data as crucial to the curriculum developer.

The program developer, unfortunately, has to attend to the demands of all of the persons in the hierarchy and provide data to meet their needs. The reasons for measuring must be answered early in program development, because the nature of the data to be collected shapes the design, the instrumentation and the interpretations that can be made (Messick, 1975). Each "audience" will be considered in turn.

The Child. First of all, measurement of the child's current level of functioning is the base upon which instruction should begin. Instruction which is too low or too high does not lead to change. Not knowing the level of the child's skill acquisition seriously diminishes the probability that learning will occur. Thus, assessment not only provides the level of initial instruction but, in many cases, indicates what should be the content of instruction.

Secondly, observing a particular child in context is the major way to determine the meaning of his or her behavior. This is particularly dramatic in language development as deLaguna (1927) demonstrated long ago and Bloom and Lahey (1978) reconfirmed more recently. For example, the two-word utterance "mommies spoon" can indicate possession (it is mommie's spoon), location (the spoon is on the table), rejection (refusing to eat off the spoon, it's not mine, it's mommie's), action (give me some more of what mommie has on her plate). The only way in which the meaning of these utterances can be determined is to observe the child and note his nonverbal actions. Even then the observer may have to repeat the full sentence for the child, to be sure the interpretation of the utterances is accurate. Understanding a child's communicative attempts is critical in facilitating his language acquisition (Snow and Ferguson, 1977). The failure

of a caregiver to try to understand the child's verbal or nonverbal communication may result in the child becoming upset and frustrated and failing to learn as rapidly as possible (Nelson, 1973). Third, data about a child's progress encourages him or her to improve. Knowing how well one can do is generally agreed upon as having a positive effect on future performance. Behavioral strategies studies have demonstrated that positive feedback helps the learner maintain a skill and provides the learner with a stimulus to remain in the learning setting.

Finally, the focus on the child tends to maximize the importance of the child as a person and his or her desirableness, not as a unit of study or a statistic to be used in the outcome data, but as the reason for the program's existence.

Adults as Teachers. House (1976) perceives major changes in the theoretical writings on evaluation strategies. "The theorists have made evaluation more democratic by including increasing numbers of people in the evaluation plan as both sources of data and as audiences" (p. 81). As House (1976) notes, being more democratic did not necessarily mean that the evaluations were necessarily more just (as in justice). He insists on the basic liberties of parents, teachers, and students, "including self-esteem not be infringed upon" (p. 97). Given the acceptance of this notion, data from pupil progress helps the teacher and parent measure their own impact on the child and enhance or maintain their own self-esteem. When data are made available to these individuals, I would argue, they are less likely to be used as evidence of teacher-parent-aide failure. Blaming the teacher (or parent) for the child's lack of progress or the program's failure is an unfortunate diversion from the task of analyzing how to modify the program so that progress can be made.

The "meaning" of data on child progress can probably best be interpreted by the persons closest to the child. A teacher or parent may see in progress something that development alone could account for or see the "breakthroughs" the experiences brought about.

The provision of progress data to the parent, for example, can help offset the negative impact the handicapped child's impairment so often has on the family by helping the parent see that their child can learn. Progress data encourages acknowledging the child's impairment, a major step in accepting the fact that the child is impaired, and setting reasonable goals for him.

The adults who work with seriously and profoundly impaired children are subject to discouragement. Madeline Appel, Chief of the Child Development Section at Roosevelt Hospital in New York City, systematically provides her staff data about child progress in a unique way. When she notes her staff's nonverbal or verbal despair in the progress of a profoundly handicapped child, she has the staff view the intake videotape. It demonstrates where the child was when he/she entered the program. Usually, the dramatic progress as viewed on the videotape is enough to highlight how effective the staff has been.

Basically, sharing child progress with the adults who share responsibility for the child is not only a human value which, in our rush to treat the child, may be shelved, it is also an important strategy for improving the child's program. Data provided to the teacher or parent encourages them by demonstrating their impact on the child and thereby facilitates future services to the child.

Of next importance, data on child progress is critical for the teacher in assessing the effectiveness of his or her techniques and the level of instructional intervention. The lack of specific knowledge of intervention results is as if we were hitting golf balls but never seeing where they land. This is not to say techniques per se are improved by knowledge of results alone. McKeachie (1976) notes that, when correction is needed, knowledge of results is most effective when accompanied by alternatives for action and the opportunity to take action. Thus, measuring pupil progress provides for the first step in program improvement.

Many professionals other than teachers need data on children's progress so that they will be able to improve their own techniques. To highlight how data may be shared to improve a program, two will be considered here, the curriculum developer and the physician. Curriculum designers need the data on children to be able to assess the effectiveness of their program. Historically, evaluation has served this purpose (Tyler, 1949). The teacher often performs the role of curriculum developer. In many programs, professional curriculum developers are not available. However, when a curriculum developer is available, the teacher's impressions along with the child data will do much to assist him or her in improving the curriculum's design.

In the same manner, the physician who referred the child should be provided data which clarifies his or her judgment. Assuming that physicians or other agency personnel can improve their referrals without feedback is often a gap in a program's operation. Sharing a delayed child's progress with physicians may encourage more early referrals rather than "let's wait and see" attitudes. Joe Bleiberg in Boston (personal communication, 1978) shared progress and lack of progress data with physicians in order to enlist their aid in reducing the dosages of drugs administered to the children in order to have more responsive students. Current efforts in training pediatricians about handicapped children stress the importance of providing child progress data so that the pediatrician understands how much an impaired child can learn (Guralnick, Richardson and Kutner, 1979; Simeonsson, Kenner and Walker, 1976).

Service Providers, Advocacy, and Professional Groups. At times we forget to inform our colleagues of the success of our programs as measured by our data. As one very behaviorally oriented student informed me in class, "the field has enough miracles stories. Now give us some data." I am not sure we ever have enough miracle stories. However, miracle stories based on a description of the intervention program and accompanied by the child's progress data are very impressive. When data are strong enough, they can and should be used for knowledge production through professional journals and through sharing

with your own colleagues.

Not all professionals in the field agree about the utility of early intervention. As Lazar et al. (1977) point out, some of those who speak most strongly against the need for early intervention are in our most prestigious university, Harvard. "It is from Harvard that we hear that intelligence is primarily genetic (Hernstein); . . . that early trauma will be 'outgrown' anyway, without the taxpayer's investments (Kagan); and that in any event, if a poor kid makes it, it's a matter of luck (Jencks)" (p. 3). These opinions have had wide dissemination, and if they are to be countered it must be done by providing data that show the effectiveness of intervention and stress the long-term effects of such intervention.

Advocacy groups are in constant need of communication to help decision makers allocate resources to children. Because a person or group is an advocate is not reason to ignore them and to focus attention on unknown groups. (A case study is data. An anecdote is data. And these data should be shared.) Advocates probably more than any other group need new positive information to keep their efforts going. This is a simple but profound truth which Dostoyevsky (1945) stated in *The Brothers Karamozov:* Miracles are for those who believe—it fulfills their hopes and desires, it adds richness to their beliefs.

The agency in which the project is housed is in need of data as to how well the students are progressing. A private agency has its board of directors. A school district has constant pressures and demands for the allocation of resources to many special programs: remedial reading, foreign language at the elementary school, consultants in speciality subjects, and on and on. Preschool education is not well fixed in the public school image as yet, and constant attention needs to be given to demonstrating the validity of your efforts.

In essence, the demand for data is related to the general scarcity of resources and the demand for evidence that our efforts make a difference. I have elsewhere stressed the need for attention to *improving* the program rather than *proving* the program (Anastasiow and Hanes, 1976). However, continuous feedback of progress data into the system does not resolve the dilemma of the pressure of accountability as measured by some type of summative data. Regardless of what orientation the evaluator and staff take on the issue of the worth of the program, at some point in time, program directors and their staffs will be pressed to demonstrate that the expenditure of the resources makes a difference and that the difference can be seen in the children. The question, as Gallagher, Surles and Hayes (1973) point out, may have been asked by the funding officer, but it may well have come through the federal labyrinth via congressmen and may have originated in the White House. It is argued here that well designed and collected progress data can be used as summative data when the need occurs.

The policy maker may sit on the board of directors of a private foundation, be a member of the Cabinet or Congress or a president of a company of one of the many interlocking board of directorships. These people are most impressed by what House (1976) refers to as the utilitarian moral value system and that value system demands accountability data. Progress data which is anecdotal may not have the impact you desire on the policy makers. Progress data put in the form of summative data (how many children made how many gains in what areas) will do much to relieve the pressure. It is safe to assume that the lack of such data may seriously undermine your efforts.

Summary

Evaluation theorists disagree on the central or primary target for your data. However, the premise of this chapter is that, for the project director, all the various audiences (children, parents, teachers, decision and policy makers) are at some time the target—frequently, for very different reasons. The bias of this chapter is that the most fundamental recipient of the data should be the child, so that his growth and development are facilitated. Development is a product of biology (maturation), the activity of the learner, and the learner's environment (including teachers and parents as well as objects). The strategy for collecting progress data in the most meaningful sense must be *developmental* in orientation and complex in design.

In the first year of the project, one may measure pupil progress in order to plan the level of instruction, select and modify instructional techniques and curriculum, and give necessary feedback to children, teachers, parents, and aides. The second year data may be used for all of these purposes and be even more helpful in refining, modifying, and adapting the instructional strategies and curriculum. It may not be until the third year that the program is robust enough to allow you to validate the fact it is having the intended impact on children or parents. When that is the case, child data are invaluable because they provide the kinds of data most desired by decision and policy makers. Say what you will, in the final analysis they will want to know: Did the children improve and can you demonstrate that your program's effect was related to the improvement?

References

Anastasiow, N. J., and Hanes, M. L. "A Case Study of Evaluating Follow Through Programs." In N. J. Anastasiow and M. L. Hanes (Eds.), *Viewpoints: Establishing and Evaluating Intervention Programs* (Bloomington, IN: Bulletin of the School of Education, Indiana University) 52 (1976): 37-61.

Anastasiow, N. J., and Majer, K. S. "Evaluating Questions for Teachers in New Programs." *Viewpoints* (Bloomington, In: Bulletin of the School of Education, Indiana University) 49 (1973): 101-110.

Blieberg, J. Personal communication, 1978.

Bloom, L., and Lahey, M. *Language Development and Language Disorders.* New York: John Wiley & Sons, 1978.

Brych, A. S., Meisels, S. J., and Markowitz, M. T. "Assessing the Effectiveness of Open Classrooms on Children with Special Needs." In S. J. Meisels (Ed.), *Special Education and Development: Perspectives on Young Children with Special Needs.* Baltimore, MD: University Park Press, 1979.

Cronbach, L. J. "Beyond the Two Disciplines of Scientific Psychology." *American Psychologist* 30 (1975): 115-127.

Darrell, L. *Justice.* New York: E. P. Dutton, 1957.

deLaguna, G. A. *Speech: Its Function and Development.* New Haven: Yale University Press, 1927.

Dostoyevsky, F. *The Brothers Karamozov.* New York: Modern Library, 1945.

Gallagher, J., Surles, R., and Hayes, A. *Program Planning and Evaluation.* Chapel Hill, NC: Technical Assistance Development System, 1973.

Glass, G. V. "A Paradox About Excellence of Schools and the People in Them." *Educational Researcher* 4 (1975): 9-13.

Guba, E. G. "The Failure of Educational Evaluation." *Educational Technology* 9 (1969): 29-38.

Guralnick, M. J. "A Research-Service Model for Support of Handicapped Children." *Exceptional Children* 39 (1973): 277-282.

Guralnick, M. J., Richardson, H. B., and Kutner, D. "Pediatric Education and The Development of Exceptional Children." Paper presented at Washington, D. C., 1979.

House, E. R. "Justice in Evaluation." In G. V. Glass (Ed.), *Evaluation Studies Review Annual* (Vol. 1). Beverly Hills, CA: Sage Publications, 1976.

Katz, L. G. "Early Childhood Programs and Ideological Disputes." *Educational Forum* 10 (1975): 267-271.

Kluger, R. *Simple Justice.* New York: Knopf, 1976.

Kohlberg, L., and Mayer, R. "Development as the Aim of Education." *Harvard Educational Review* 42 (1972): 449-496.

Lazar, I., et al. *The Persistence of Preschool Effects. A Long-Term Follow-up of 14 Infant and Preschool Experiments.* Washington, DC: U. S. Government Printing Office, 1977.

McCall, R. B. "Challenges to a Science of Developmental Psychology." *Child Development* 48 (1977): 333-344.

McKeachie, W. J. "Psychology in America's Bicentennial Year." *American Psychologist* 31 (1976): 819-833.

Magoon, A. J. "Constructivist Research." *Review of Education* 49 (1977): 651-694.

Messick, S. "The Standard Problem. Meaning and Values in Measurement and Evaluation." *American Psychologist* 30 (1975): 955-966.

Nelson, K. "Structure and Strategy in Learning to Talk." *Monographs of the Society for Research in Child Development* 38 (1973): (1-2, Serial No. 149) 135.

Riegel, K. F. "Influence of Economic and Political Ideologies on the Development of Developmental Psychology." *Psychological Bulletin* 78 (1978): 129-141.

Rudel, R. G. "Neuroplasticity: Implications for Development and Education." In J. S. Chall and A. F. Mirsky (Eds.), *Education and the Brain.* Chicago: University of Chicago Press, 1978.

Scriven, M. "Maximizing the Power of Causal Investigations: The Modus Operandi Method." In G. V. Glass (Ed.), *Evaluation Studies Review Annual* (Vol. 1). Beverly Hills, CA.: Sage Publications, 1976.

Simeonsson, R. J., Kenner, W., and Walker, L. "Child Development and Disability: Competency-based Clerkship." *Journal of Medical Education* 51 (1976): 578-581.

Snow, C. E., and Ferguson, C. A. *Talking to Children: Language Input and Acquisition.* Cambridge, MA: Cambridge University Press, 1977.

Stake, R. E. "The Countenance of Education Evaluation." *Teachers College Record* 68 (1967): 523-540.

Stengel, A. "Evaluation Heuristics." Research Paper. Bloomington, IN: Institute for Child Study, 1978.

Takanishi, R. "Evaluation of Early Childhood Programs: Toward a Developmental Perspective." In L. G. Katz (Ed.), *Current Topics in Early Childhood Education* (Vol. 2). Los Angeles: University of California, 1978.

Tyler, R. *Basic Principles of Curriculum and Instruction.* Chicago: The University of Chicago Press, 1949.

Tyler, L. L. "Curriculum Evaluation and Persons." *Educational Leadership* 35 (1978): 275-279.

Werner, E. E., and Smith, R. S. *Kauai's Children Come of Age.* Honolulu, HI: University Press of Hawaii, 1977.

SECTION 2: INSTRUMENT SELECTION

Editor's Notes

The staff of a project measures children in several ways and makes numerous decisions that affect how they are measured. Sometimes it is difficult to sort out those decisions, how they affect each other, and how they relate to the overall program for children. The chapters in this section are organized around a conceptual framework that has two dimensions: the types of children being served and the purposes for which they are being measured.

Handicapped children can be separated into types in several ways, e.g., by handicapping condition, by age, by degree of impairment, etc. In this section children are separated into three general categories: the mild to moderately handicapped preschool child (ages 3-5), the mild to moderately handicapped or "at risk" infant (ages 0-3), and the severely and profoundly handicapped child (ages 0-5). Four general purposes for measuring these children are identified: *screening, diagnosis, assessment,* and *evaluation.*

These four terms are used quite often in early childhood special education, and, for the most part, interchangeably. Here, each term has a distinct meaning that distinguishes it from the others: *Screening* is measurement that seeks either to raise or allay suspicion of a handicap and leads to a decision about whether further diagnostic examination is needed. *Diagnosis* seeks to confirm whether or not a handicap(s) exists and to gather information needed in order to decide on appropriate program placements. *Assessment* occurs after placement. It seeks to examine strengths and weaknesses and leads to planning educational and other interventions. *Evaluation* is measurement that seeks to determine the effectiveness of the intervention program by examining the progress of children. It leads to decisions about program improvement and to judgments about the value of the program. One distinction that separates *evaluation* from the other three terms is its dual focus on both the child and the program. For *screening, diagnosis,* and *assessment,* the primary object of inquiry is the child, although the family and other environmental factors are usually examined also. However, in *evaluation* the primary object of inquiry is the intervention program with evidence of child progress becoming the means by which one measures program effectiveness.

In terms of the purposes for measurement and the types of children being measured, Section II examines some of the issues that early childhood special educators must face. Each chapter provides information that should be taken into account when making decisions about measurement and suggests strategies for selecting the instruments and procedures that will best meet the needs of a particular program. They also identify some of the instruments and procedures currently being used.

Kemper and Frankenburg lead off the section with a comparative discussion of *screening, diagnosis,* and *assessment.* Each is described in terms of its purpose, who conducts the measurement, and how the results are used. They also discuss reliability and validity and provide tables with information on several screening, diagnostic, and assessment tests.

In Chapter 3, Harbin examines assessment of the mild to moderately handicapped preschooler (ages 3-5). The thesis of the chapter is that the assessment devices chosen should be compatible with the philosophy and goals inherent in a project's curriculum model, as well as with the characteristics of the children being served, if they are to be useful in planning educational programs. Six curriculum models for preschool children, including their approaches to assessment, are described; and examples of assessment devices for each model are given. The author also provides special information on assessment devices in three curriculum areas: language, social/emotional, and perceptual development.

In Chapter 4, Sheehan focuses on selecting instruments for program evaluation. While there are no tests specifically designed for this purpose, he suggests that tests and procedures designed for other tasks—e.g., screening, diagnosis, and assessment—can be considered in evaluating programs. He also addresses some of the problems and limitations of such instruments when used in this way.

In Chapter 5, Kopp begins by presenting a broad perspective of the very young child's development. She highlights both biological and environmental influences, especially the critical role of the "caregiver." She then argues that parents should be the central focus of an infant program's initial intervention, and, consequently, the subject of assessment and program evaluation efforts. Once a good working relationship between parents and program exists, attention can be turned to the infant. While a developmental perspective is essential to measuring child progress, Kopp cites problems and limitations in our current methods and concludes by identifying some additional areas of measurement that the field should explore in order to obtain meaningful documentation of child progress.

The *assessment* of "at risk" infants and the *evaluation* of infant intervention programs are the topics of Chapter 6. Badger stresses the need to assess both the mother and the infant in order to plan an intervention program and to evaluate program effectiveness. She identifies three instruments used to examine the mother's caretaking ability and three instruments used to measure infant development.

Chapters 7 and 8 deal with assessment and evaluation of the severely and profoundly handicapped child. Robinson, in Chapter 7, lays out the problems inherent in using norm-referenced developmental tests with these children. Using the *Bayley Scales* as an illustration, she describes several ways norm-referenced instruments are not appropriate. She then offers specific criteria for judging the utility of an assessment procedure for the severely handicapped, and employs the *Uzgiris-Hunt* scales to illustrate those criteria. She also points out that in evaluating the effectiveness of intervention with these children, "What does effective mean?" is the question that must be answered first.

Assessment of communication in the severely handicapped is the focus of Chapter 8. DuBose presents a description of communication and expressive language development and relates it to assessment. She also identifies several instruments and curricula for this area.

CHAPTER 2

Screening, Diagnosis, and Assessment: How do these types of measurement differ?

by Melinda B. Kemper and William K. Frankenburg

Introduction

The purposes of this chapter are to discuss 1) the differences between screening, diagnosis, and assessment as these activities relate to the identification of children with handicapping conditions and 2) the planning of appropriate educational activities for such children. Screening identifies children who may have a handicapping condition and who need further evaluation. Diagnosis confirms or disconfirms the existence of a handicap suspected on screening. Assessment determines a child's current level of functioning and aids in the development of individual educational plans. Tests which are designed to implement screening, diagnosis, and assessment are different in content and purpose. They are intended for different groups of children and require different levels of expertise for proper administration. (See Table 1).

If the goals of early identification and intervention activities are to be achieved in a cost-efficient manner, people must understand the differences between the three types of tests. Screening, diagnosis, and assessment are treated separately in the following sections. Each activity is defined and its purposes explained. The nature of the tests used for each is also described, and information is provided on who gives the test, who receives it, and when the test should be used.

Table 1

Types of Testing Procedures

Type of Testing	Purpose of Testing	Nature of Tests and Procedures	Who Gives the Test	Who is Tested
SCREENING	To identify children who are in need of further evaluation	Tests which are quick, simple, economical, and valid	Paraprofessionals or volunteers	All children in target population
DIAGNOSIS	To determine whether problems exists To determine nature and extent of problem To determine etiology and prognosis if possible To determine appropriate program placement	Various procedures including but not limited to norm-referenced tests with predictive validity	Professionals (Pediatricians, Psychologists, Speech Pathologists, Occupational or Physical Therapists, and others)	Children who have been identified as suspect on screening
ASSESSMENT	To assess current specific skills To establish short- and long-term goals and objectives To monitor progress	Developmentally sequenced criterion-referenced measures	Professionals (Special Education and Classroom teachers and others)	Children who have been diagnosed as having a problem and who have been placed in treatment

Screening

Screening is defined as the presumptive identification of unrecognized disease or defects by the application of tests, examinations or other procedures which can be applied rapidly. Screening tests are designed to be administered to large numbers of asymptomatic individuals in order to select those who should receive additional evaluations. Screening tests do not determine *who* has problems; they identify individuals who have a high *probability* of having problems and who require further evaluation to determine if such problems actually exist. The purpose of screening is to *move up* the time of diagnosis and treatment. For instance, without mass screening efforts, handicapped children, particularly those with milder handicaps, would often not be identified until after they had entered school. By that time, several years would have elapsed during which treatment might have been more effective. Both behavior and family problems might have developed secondarily to the child's handicapping condition.

Since screening tests are designed to be given to large groups of individuals, they are generally quick, simple, and economical to administer. They have a limited number of items and are usually scored as suspect or not suspect or as normal, questionable, or abnormal. They are generally given by paraprofessionals or volunteers since the use of professionals would be extremely costly and since the available professionals are likely to be required for diagnosis and treatment.

Since the purpose of screening is to identify children who need further evaluation, tests must be used which compare a child to the general population. Therefore, screening tests should be standardized. A standardized test is one that has been given to a well-defined group (standardization sample) and for which norms have been established. Administration and scoring proce-

TABLE II

SCREENING TESTS

TEST	AGE RANGE	STANDARDIZATION	RELIABILITY	VALIDITY	TIME	COMMENTS
Denver Developmental Screening Test (DDST)	0-6 years	1,036 children representative of Denver population	Test-retest .97; Inter-observer .90	Identification of concurrent retarded development 85-100%; Predictive validity of Abnormal score for for later school problems = 90%	15-20 mins.	Items cover gross motor, fine motor-adaptive, language, and personal-social skills. May be administered by trained paraprofessional, scored as Normal, Questionable, Abnormal and Untestable. A short form of the DDST has recently been authorized as a prescreen for the full DDST.
Prescreening Developmental Questionnaire (PDQ)	3 months to 6 years	none as such; item placement based on DDST standardization sample	?	93% item agreement with DDST	5 mins.	Parent answered questionnaire that covers same areas of development covered by DDST. Designed as a prescreen for the DDST. May be given by paraprofessional. Less accurate when used by parents with low education levels.
Developmental Screening Inventory (DSI)	1-18 months	none; item placement based on Gesell	Interobserver .85 to .90 within 10 points of score assigned	DSI and Gesell: .38-.88; Prediction of later abnormality: .81	20-30 mins.	Items cover gross motor, fine motor, adaptive, language, and personal-social skills. Designed for nurses and physicians, it can be used by others who are experienced in infant evaluation.
Developmental Indicators for the Assessment of Learning (DIAL)	2½ to 5½ years	4,356 children stratified sample	Interobserver between two authors of test was reported as high	Distinguished between extreme groups (upper 90% and lower 10%); after 1 year concurrent agreement between DIAL and various other tests: .85	20-30 mins.	Items cover gross motor, fine motor, concepts, and communications. Designed to be administered by one professional as coordinator and four paraprofessionals as operators to administer items in each area tested. Children are tested individually, but several may be involved in testing at one time.
Quick Test	2 years to adult	458 children 2 years to 12th grade	interform reliability for preschoolers .78 to .92	with Full Range Picture Vocabulary Test .62-.79	5-10 mins.	Quick screen of verbal-perceptual intelligence. May be administered by paraprofessionals

dures must also be standardized in the sense that they are clearly specified and made uniform so that a test may be given in a consistent manner.

It is important that screening tests be reliable or consistent in their results. If test results are inconsistent, it is unlikely that they will be accurate. *Interobserver reliability* means that two people looking at the same test behavior or performance would derive similar test scores. *Test-retest reliability* is the agreement of test results when the same test is administered on two separate occasions to the same individual. *Reliability* estimates indicate the extent to which a test score is likely to be free of chance inaccuracies or variable errors. It is important that paraprofessionals and volunteers be carefully trained in the administration and scoring of screening tests, since even a good test will yield meaningless results if it is not given according to its standardized procedures. Some screening test developers provide programmed instructions so these standardized procedures may be learned close to home at minimal cost.

Screening tests must be *valid*. They must measure accurately what they purport to measure. Such accuracy can be established by a concurrent agreement with a standard diagnostic test or with a diagnosis established independently by other examiners. Another type of validity is the ability of a screening test to *predict* the occurrence of a problem at a later date. The agreement between screening and diagnostic findings can be classified in several different ways. Two of the most useful ways are sensitivity and predictive value of a positive. *Sensitivity* is the percentage of the handicapped who are identified as suspect on screening. Ideally, one would like to see sensitivity equal 100 percent, meaning that there were no missed cases. In practice, this almost never happens. Furthermore, a screening test that is highly sensitive in terms of not missing a handicapped child could also identify many children as suspect who really don't have a handicap. These are called overreferrals. On the other hand, there would be few overreferrals if the *predictive value of a positive* were also high. *Predictive value of a positive* is the percentage of those suspect on screening who actually turn out to have a handicap on diagnosis.

Screening tests are intended not only to be administered to large groups of individuals but to be periodically readministered to large groups. Periodic rescreening is especially important when screening very young children for handicapping conditions since problems can develop at any age.

Table II gives a brief summary of the characteristics of some of the more commonly used developmental screening tests.

Diagnosis

Diagnosis confirms or disconfirms the existence of a suspected handicapped condition. Children who are suspect on screening are referred for diagnostic evaluations. Whereas screening tests sample a limited number of skills in one or more areas of development, diagnostic tests provide an in-depth look at one or two areas of development. Diagnostic tests are more complex, more time consuming, and more expensive to administer than screening tests, and they must be administered by professional specialists who have certain prerequisite qualifications and training.

There are several issues that a diagnostic evaluation must address. First of all, the existence of a handicapping condition must be ruled in or out. If a problem does exist, it is necessary to determine its *nature* and *extent*. Is it a minor variation which is unlikely to have long-term significance, or is it a major problem that if not treated will adversely affect a child's future learning? Next, it is useful to determine the *etiology* or cause of the problem. Knowing the cause may have implications for treatment and for prognosis, with and without treatment. For instance, it is helpful to know if a language delay is due to a lack of environmental stimulation or to a significant hearing loss or to some other condition. Sometimes etiology is unclear. *Prognosis* is another aspect of the diagnostic procedure. Often prognosis is a matter of making an educated guess based on known outcomes of similar cases. But since there are seldom any two cases exactly alike, prognoses are usually phrased as statements of probability. Prognosis is likely to be affected by the type of treatment a child receives. And a child's response to treatment will clearly affect prognosis.

The tests used in a diagnostic evaluation depend upon the nature of the handicapping condition in question. For a child who was suspected of being mentally retarded or developmentally delayed, one would need to obtain a complete developmental history and perform a developmental or intellectual evaluation with a standardized, norm-referenced test. If the child's developmental milestones and test performance appeared to be within normal limits and no other problems were suspected, that might be all that was necessary; however, if the child appeared to be delayed, additional evaluations would be necessary to clarify the nature and extent of the problem and to investigate the possible etiology or cause. Information which would help to clarify the nature and extent of the problem would come from speech and language, occupational therapy, physical therapy, and social-emotional evaluations. Etiology would be investigated by obtaining a complete medical history, a thorough physical exam, a neurologic exam, a hearing test, and possibly a vision test, a chromosome analysis, metabolic screening, an EEG, or a CAT scan, depending upon the nature of the child's disability. This distinction between evaluating the nature of the problem and investigating its etiology is not clear-cut. For instance, a developmental delay could be rooted in emotional difficulties. The foregoing illustrates the potential depth and complexity of a diagnostic evaluation.

Just as standardization, reliability, and validity are critical in the selection of screening tests to be used by paraprofessionals, these test characteristics are important considerations in the selection of diagnostic tests; however, there are occasions when well-trained, highly skilled professionals choose tests that have not been demonstrated to meet these criteria. There are some areas of development for which we do not have well-standardi-

zed diagnostic tests, and there are some children whose ages or specific disabilities make standardized testing difficult. If a child's level of functioning is near or below the limits of the norms provided for a particular test, only a small sample of behavior will be obtained, rendering the use of norms and resulting conclusions less meaningful. When specific deficits are known, such as poor vision or poor verbal skills, the selection of assessment procedures should be accommodated to a child's special needs. An experienced clinician may use a battery of tests or subtests plus informal testing and observation in order to assess a child's level of functioning in a given area of development. The needed variations in the testing procedures are the responsibility of the examiner. Testing handicapped infants and young children can be very difficult.

Obviously, it is extremely important that diagnostic evaluations be performed by professionals who have a thorough understanding of both normal development and developmental disabilities as well as specialized clinical training. Diagnostic procedures are no better than the skills of the person using them, and test results can be very damaging when misinterpreted or poorly understood.

In programs like Child Find, diagnostic tests are only administered to children who are suspect on screening or who are suspect due to concerns of parents or professionals who have been involved previously with the child or the family. The diagnostic evaluation by various professionals ends in an interdisciplinary staff conference during which various findings are integrated to arrive at a comprehensive diagnosis, to provide information on etiology and prognosis, and to determine appropriate program placement.

Table III gives an overview of some of the more common diagnostic tests and procedures which are used to clarify the nature and extent of various handicapping conditions.

Assessment

Educational assessment is designed to measure a child's existing skills in a variety of different areas such as cognitive, language, gross motor, fine motor, social, and self-care. These assessment instruments are checklists of abilities or functions which are developmentally sequenced within each area. Tasks not yet performed or skills not yet acquired become specific instructional objectives. The purpose of such an evaluation is to aid in establishing the short- and long-term goals and objectives of a child's individualized educational plan as well as to measure progress and set new instructional objectives at regular intervals. Unlike diagnostic tests which are generally limited to one or two areas of development, assessment instruments usually cover many areas. They often contain a larger number of items and provide smaller sequential steps in each area covered in order to facilitate educational programming.

Assessment instruments are administered by the person who is most in need of the information for programming. This is generally the classroom teacher and may include speech, physical and/or occupational therapists who are working with and planning for the child. Assessments are conducted for children who have been diagnosed as having a handicapping condition and after they have been placed in a treatment program. They are usually conducted over a period of time after the child has become accustomed to the teacher and the classroom setting and after the teacher has become familiar with the child's habitual mode of behavior.

Assessment scales are numerous and are generally not standardized. Few have any documented reliability or validity data. Reliability is not generally a problem since most assessment scales are intended to be used by classroom teachers over an extended period of time. Whenever a child is deemed to perform a task adequately, then the next task in the developmental sequence becomes an objective in that child's educational plan. Validity is, of course, implied, but content validity and careful sequencing are critically important since the assessment tasks will become the basis of an educational program. The specific skills tested must be relevant to program goals.

Lack of standardization means that these assessment instruments do not have normative data. They are not designed to compare a child's performance to others of his age. They are designed to compare a child to a set of standards or criteria and, for this reason, are called criterion-referenced tests. Many of these tests have been developed by taking items from various standardized tests and placing them at age levels corresponding to the age level at which they appeared on the standardized test. This may be useful for indicating *approximate* ages at which various skills or behaviors may be expected, but one cannot make normative assumptions about items taken from other scales for several reasons:

1. These items are no longer administered in the order and context in which they were administered during standardization.
2. Administration procedures are likely to vary, since the instructions in criterion-referenced tests are limited compared to the extensive instructions provided in the tests from which the items were borrowed.
3. Scoring procedures are likely to vary for the same reason as administration procedures.
4. Different norm-referenced tests have employed different standardization samples, different item-administration instructions, and different scoring criteria, and the age placement of apparently similar items will vary from test to test.
5. Different test developers place items differently with regard to the percent of children of a given age who pass those items. On some tests, items are placed at a given age level when 50 percent of the children that age can perform the item. On other tests, an item is not placed at a given age level unless 60-75 percent of the

children that age can perform it.

These points are mentioned to caution the potential user regarding the use of assessment tests as diagnostic instruments.

It should be noted that while assessment instruments are not likely to be useful as diagnostic tools, diagnostic tests are also generally inadequate for the purposes of daily programming. They are likely to contain fewer items than good criterion-referenced tests, and the items themselves are frequently not relevant for programming. Also, diagnostic/norm-referenced tests must be administered by trained professionals whose services are not necessarily available to classroom teachers on an ongoing basis. Nevertheless, it is important for the classroom teacher to have diagnostic information available because this information will help in the setting of long-term goals and in the determination of the most useful strategies for achieving them.

Table IV summarizes some of the more commonly used assessment scales.

Summary

It is important that the purposes of screening, diagnosis, and assessment be clearly understood so that tests designed for these purposes are used appropriately. The purpose of screening is to identify children who need further evaluation. Screening tests can be administered by trained paraprofessionals or volunteers and should be quick, simple, economical, and valid. Screening tests are given periodically to all children in the target population.

Diagnostic tests are given to children who have been identified as suspect on screening. The purpose of diagnosis is to determine whether a handicap exists, to determine the nature and extent of the handicap if it does exist, to determine etiology and prognosis if possible, and to determine appropriate program placement. Diagnostic evaluations are performed by professionals with highly specialized training in various disciplines. Diagnostic evaluations employ various procedures including standardized tests with predictive validity.

Assessment tests are given to children following program placement and at regular intervals thereafter to assess specific skills in a variety of areas. The purpose of assessment is to aid in establishing instructional objectives and to monitor progress. Assessment instruments are developmentally sequenced, criterion-referenced tests which are administered by classroom teachers and others who are responsible for setting and implementing educational objectives.

Proper use of screening, diagnostic and assessment tests ensures that the goals of early identification and treatment will be met most effectively.

TABLE III

DIAGNOSTIC TESTS: SOCIALIZATION AND SELF-HELP SKILLS

TEST	AGE RANGE	COPYRIGHT DATE	STANDARDIZATION	RELIABILITY	VALIDITY	USE
Vineland Social Maturity Scale	All ages	1965	620 subjects 0-30 years	Test-retest: .94-.99 for social age; .57 for social quotient. Parents vs. teacher reports: .92	Not established with external criterion; Limited relationship with Binet	Structured parent interview. Assessment may reveal signs of overprotection, understimulation, or emotional immaturity. Value limited by reliance on secondary sources, but sometimes useful in determining whether a child's independent functioning is within the limits of his intellectual abilities.
Informal Observation	A child's social skills may sometimes be assessed by informal observation in a variety of situations. These would include a play session with an examiner, a parent-child play interaction, and free play with peers in a preschool setting. Useful information can be obtained about a child's style of relating to others and to play materials by observing such interactions. This type of assessment would require a psychologist who had extensive experience in this area.					

TABLE III

DIAGNOSTIC TESTS: PERCEPTUAL-MOTOR SKILLS

TEST	AGE RANGE	COPYRIGHT DATE	STANDARDIZATION	RELIABILITY	VALIDITY	USE
Ayres Battery	4 to 7-10 years depending on subtest	1964 to 1969 depending on subtest	Different limited samples for different subtests	Test-retest: .01-.93	Differentiates between clinical groups	Various subtests provide potential breakdown of problems underlying functional deficits in fine motor area. Not useful with children who have CP or other severe motor problems. Used with children who have or are suspected to have visual-perceptual or visual-motor problems. Not generally useful until age 5.
Marianne Frostig Developmental Test of Visual Perception	3-8 years	1966	2,116 mid-SES children	Test-retest: .42-.80	Frostig & first grade reading achievement: .40-.50	Useful in assessing eye hand coordination and readiness for attending in a school setting
Developmental Test of Visual-Motor Integration (Beery)	2-15 years	1967	1,039 children	Test-retest: .83-.87	Beery & first grade reading achievement: .50	Often used as a screening test for prewriting skills
Bender Visual Motor Gestalt Test	4 years to adult	1974	975 children 5-11 years, not geographically representative	Interobserver: .88-.96 Test-retest: .60-.66	Bender & Binet or WISC: .48-.79 Bender & first grade reading achievement: .37-.67	Test is sometimes given by psychologist - rarely by occupational therapists. Developmental levels can be scored reliability but neuropsychological and emotional indicators are less reliable. Not useful below 6 years.
Occupational Therapy Evaluation	While the Ayres Battery, the Frostig, and/or the Beery are sometimes used when fine motor or perceptual motor deficits are suspected or documented, it is important that these tests be used in conjunction with an assessment of a child's primitive reflexes, postural reactions, muscle tone, sensory awareness, joint stability, and coordination (see comments under Gross Motor Development).					

TABLE III

DIAGNOSTIC TESTS: GROSS MOTOR SKILLS

TEST	AGE RANGE	COPYRIGHT DATE	STANDARDIZATION	RELIABILITY	VALIDITY	USE
Bayley Scales of Infant Development: Motor Scale	2-30 mos	1969	1,262 infants	Inter-rater: .93 Split-half: .63-.92 Test-retest: .75		Assess gross motor and some fine motor skills. Test of choice for children in age range. Sometimes useful for assessing developmental level of motorically retarded older children. Assesses functioning level but not quality of movement.
Physical Therapy Evaluation	No standardized test for this type of evaluation. Where motor problems are suspected, it is important that a child's reflexes, reaction patterns, muscle tone, coordination, sensory awareness, joint stability, and functional level be evaluated by a physical or occupational therapist who is experienced with infants and young children.					

TABLE III

DIAGNOSTIC TESTS: COGNITIVE SKILLS

TEST	AGE RANGE	COPYRIGHT DATE	STANDARDIZATION	RELIABILITY	VALIDITY	USE
Bayley Scales of Infant Development: Mental Scale	2-30 mos.	1969	1,262 infants; representative sample	Inter-rater: .89 Split-half: .81-.93 Test-retest: .76	Bayley & Binet 24-30 months levels: .57 Infants scoring >2SD below the mean are likely to have continuing delays.	Assess attentional, fine motor, problem solving, language, and other areas of early cognitive development. Test of choice for children 2-30 months of age. Sometimes useful for assessing developmental level of older retarded children.
Stanford-Binet Intelligence Scale (Form L-M)	2 years to adult	1972	2,351 children, 2-18 years; representative sample	Standard error of measurement: ±5	Many studies available. Predicts scholastic achievement.	Primarily verbal items, with some performance items for preschool years. Frequent test of choice for children 2½ to 5 years, unless child is expected to function below 2-year level.
Wechsler Pre-school and Primary Scale of Intelligence (WPPSI)	4-6½ years	1967	1,200 children, representative sample of 1960 census	Split-half: .71-.96 Test-retest: .60-.93	WPPSI & Binet: .32-.76. Predicts scholastic achievement although less well than Binet.	Frequent test of choice for children in this age range, but limited usefulness for those who are retarded. Provides separate verbal and performance scores.
McCarthy Scales of Children's Abilities (MSCA)	2½-8½ years	1972	1,032 children; representative sample of 1970 census	Split-half and test-retest: .79-.93	MSCA & Binet, MSCA & WPPSI: .62-.81 Predicts to scores on Metropolitan Achievement Tests with normal population and much less well when child is delayed.	Considered one of least "culturally biased" tests. Has six subscales and is most useful with children who are not retarded but who have mild learning disabilities.
Arthur Adaptation of Leiter International Performance Scale	2-9 years	1950	289 mid-SES 3-7 year olds	Split-half: .91-.94	Leiter-Binet: .64-.81	Useful with children who have auditory deficits, speech and language problems or motor handicaps. Assessment limited to non-verbal skills
The Extended Merrill-Palmer Scale	3-5 years	1978	1,242 white children	Not available in manual	Not established with outside criterion	Based on structure of intellect model. Test is too new for practical usefulness to be established.

TABLE III

DIAGNOSTIC TESTS: SPEECH AND LANGUAGE SKILLS

TEST	AGE RANGE	COPYRIGHT DATE	STANDARDIZATION	RELIABILITY	VALIDITY	USE
Sequenced Inventory of Communication Development (SICD)	4 mos. to 4 years	1975	Limited sample	Inter-rater: .90-1.00 Test-retest: .88-.99	SICD & PPVT: .76-.81	Frequent test of choice for children in age range. Variety of subtests measuring receptive and expressive language development.
Peabody Picture Vocabulary Test (PPVT)	2½ to 18 years	1965	4,012 white children from around Nashville	Parallel-form: .67-.84	Limited agreement with Binet & WISC	Receptive vocabulary. Useful with children with expressive deficits. Should be used only in conjunction with other tests.
Receptive-Expressive Emergent Language Scale (REEL)	0-3 years	1971	50 normal infants from linguistically enriched environments	Test-retest: .71	Not established with external criterion	Structured parent interview. Useful with children who are silent during testing or who are otherwise difficult to test. Also useful for older children who are severely delayed.
Test of Auditory Comprehension of Language (TACL)	3-6 years	1973	200 mid-SES children	Test-retest: .94	TACL & I.Q. for retarded children: .80	Receptive syntax and morphology. Useful as part of a test battery.
Illinois Test of Psycholinguistic Abilities (ITPA)	2-10 years	1960	962 children from Illinois; not representative	Test-retest: .28-.93 Interobserver: .94-1.00	Not established with external criterion	Not generally considered an adequate measure of language development when used alone, but often used as part of a test battery.
Developmental Sentence Analysis	2-7 years	1974	200 mid-SES white children	Split-half: .73	Not established with external criterion	Method of assessing recorded language samples for expressive syntax and morphology. Sometimes used as part of a language test battery. Norms appropriate only for children from mono-lingual homes who are learning standard American-English grammar.
Photo Articulation Test	3-12 years	1969	648 white mid-SES children	Test-retest: .99	PAT & Bryngelson-Glaspe: .97 PAT & Templin-Darley: .82	Diagnostic articulation test.
Goldman-Fristoe Test of Articulation	2 years to adult	1972	Not standardized for preschool ages	Test-retest: .86-.95 Inter-rater: .88-.92	Not established with outside criterion	May be more interesting for young children than Photo Articulation Test
Oral Peripheral Exam	All ages	This is a general procedure for which there is no specific formal test. This assessment involves the examination of a child's oral structure, oral motor control, and oral sensory awareness in order to assess the possibility that problems in any of these areas might contribute to an existing expressive language delay.				
Audiometric Exam	All ages	The method used as well as its reliability and validity vary with the age of the child being examined. Auditory screening is likely to be a part of a comprehensive screening program, and an audiometric exam will be part of the follow-up for suspect auditory screening results. The results of the audiometric exam, if performed by someone other than the speech pathologist, should be made available to this individual since information on hearing acuity is essential to a complete assesment of speech and language problems.				
Goldman-Fristoe-Woodcock Auditory Skills Test Battery	3 years to adult	1976	7,199 normal and clinical subjects 3 years to adult	Split-half: .77-.92	Differentiates between clinical groups	Auditory discrimination, selective attention, memory and sound-symbol subtests used to identify specific deficiences in auditory functioning. Some or all of subtests can be used as part of a diagnostic battery.

TABLE IV

ASSESSMENT SCALES

TEST	AGE	DEVELOPMENTAL AREAS						COMMENTS
		GROSS MOTOR	FINE MOTOR	CONCEPTUAL	LANGUAGE	SOCIAL	SELF-HELP	
Behavioral Developmental Profile (Marshalltown)	0-6 years		Δ		Δ		Δ	Intented for use with handicapped and "culturally deprived" in home-based programs. Designed to be used with Marshalltown Prescription Manual Curriculum.
Callier-Azusa Scales	0-Adult		Δ		Δ	Δ	Δ	Intended for use with low functioning deaf-blind and multihandicapped individuals. Reliability data available from editor.
Carolina Developmental Profile	2-5 years	Δ	Δ	Δ	Δ	Δ		Intended for use with mildly handicapped children. Designed to be used with Developmental Task Instructional System.
Developmental Checklist (Zimmerman-Bornstein)	1-8 years	Δ	Δ		Δ	Δ	Δ	Intended for use with visually impaired and severely or profoundly retarded. Items may be passed by report if observation is not feasible.
Developmental Profile (Alpern-Boll)	0-12 years		Δ		Δ	Δ	Δ	Δ	Designed as an interview, but items can be administered if desirable or necessary. Three items per 6-month age level to 3½ years and 3 items per 1-year age level to preadolescence. It has been standardized in its interview form and both reliability and validity data are available.
Learning Accomplishment Profile (LAP)	0-6 years	Δ	Δ	Δ	Δ	Δ	Δ	There are 3 forms of LAP: the diagnostic LAP, the prescriptive LAP, and the infant LAP. Limited standardization studies have been completed with the diagnostic LAP, and both standardization and validity studies are in progress. The LAP is cross referenced to a curriculum guide, and training filmstrips are available.
Memphis Comprehensive Developmental Scale	0-5 years	Δ	Δ	Δ	Δ		Δ	Each of the motor scales has 2 items per 3-month age level. The other 3 scales have 3 items per 3-month age level. Lesson plans for Enhancing Preschool Developmental Progress is a curriculum guide for the Memphis Model.
Portage Project Checklist	0-6 years		Δ		Δ	Δ	Δ	Δ	Intended primarily for use in home-based programs. Cross referenced to the Portage Guide to Early Education Curriculum.
Preschool Attainment Record	6 month to 7 years	Δ	Δ	Δ	Δ	Δ	Δ	Designed as an interview to assess children for whom direct examination is not feasible. Observation may be used to confirm report. There is 1 item for each 6-month age span in each developmental area.
Program for Preschool Visually Impaired Children Developmental Checklist	0-5 years	Δ	Δ	Δ	Δ	Δ	Δ	Intended for use with visually impaired, but without vision subsection may be used for all disability groups. Items are placed at age level when passed by normal children.
SEEC Developmental Wheel	0-5 years		Δ		Δ	Δ	Δ	Δ	The "mini" wheel gives 12 developmental milestones in the 5 areas measured. The "maxi" wheel gives a more extensive sequencing of skills in each area.
SEED Developmental Profiles	0-4 years	Δ	Δ	Δ	Δ	Δ	Δ	Intended for use with severely handicapped children. Items for the 1st year are given in 4-week increments, and then in 3-month increments thereafter.
TARC Assessment System	3-16 years		Δ			Δ	Δ	Δ	Designed as a quick assessment of severely handicapped individuals. Limited number of items.
Valett Developmental Survey of Basic Learning Abilities	2-7 years	Δ	Δ	Δ	Δ			Intended for use with developmentally delayed mildly retarded. Covers additional areas of visual, tactile, and auditory discrimination. Highly variable number of items within each area measured.
Vision Up	0-6 years	Δ	Δ	Δ	Δ	Δ	Δ	One of the most detailed assessment scales available. The large number of items may make administration lengthy, but provides very careful sequencing for purposes of program planning, especially for the severely impaired.

Δ = indicates area(s) covered in test

CHAPTER 3
Mildly to Moderately Handicapped Preschoolers: How do you select child assessment instruments?
by Gloria Harbin

People often ask, "What is the best device for assessing young handicapped children?" The answer to this question is "There is no one single device that will serve the needs of all programs that assess these children." In fact, several devices are usually required within individual programs; one device rarely provides enough information on a child for programming. The philosophical approach used by programs, the training of the professionals who conduct assessments, and the characteristics of the children who will be evaluated are all variables which must be figured into the selection of assessment instruments.

The aims of this chapter are to familiarize you with many of the devices available for the three-to-five-year-old, mildly-to-moderately handicapped group and to help you select devices that will best serve your needs. To these ends, discussion is included on (1) types of assessment, (2) characteristics of the child population, and (3) the kinds of curriculum approaches commonly in use. Special information is also provided on language, social-emotional, and perceptual assessment devices, because these are often three of the most problematic areas for educators.

Types of Assessment

There are basically four different assessment techniques: norm-referenced, criterion-referenced, observation, and Piagetian. In general, each differs in purpose, in the assumptions underlying scoring and interpretation, in who is qualified to use the technique, and in degree of applicability to programming (see Figure 1). The different techniques provide qualitatively different kinds of information. Except in a few cases, *criterion-referenced* devices and *observation* have been the most popular techniques for use in educational programming. They often give more detailed and subtle information than *norm-referenced* devices. The newer *Piagetian* devices promise to provide more information on the quality or level of the child's skill than previously possible. Most of the Piagetian devices are for children within the sensorimotor stage of development (birth - 2 years), however, and concentrate only upon cognitive development. An explanation of these four assessment types and their use can be found in Chapter 5, "Educational Assessment" (Harbin, pp. 35-52), of *Identifying Handicapped Children* (Cross and Goin, 1977). When you understand the limits of various devices, you will be more likely to use a combination of techniques which will provide more complete and less fragmented assessment information on each child.

Child Characteristics

A second major consideration in selecting educational programming devices is the characteristics of the children in your program. The level of functioning of a group of moderately handicapped children, for example, might indicate the use of a criterion-referenced device which focuses primarily on functional or daily living skills. On the other hand, for a group of mildly retarded children who may be mainstreamed into regular classes, an assessment device which contains at least an adequate portion of items on readiness skills for academics would be appropriate.

Curriculum Models

The curriculum model of your project is the third major consideration in the selection of assessment devices. Each assessment device is based upon a particular philosophical approach which influences its content. It is essential that the *type* of assessment device and its *content* be consistent with the philosophical approach of your program. Assessment is an integral part of all curriculum models; consequently, skills or behaviors which are important in your model and which you teach or emphasize with the children should be measured. It is not to your advantage, for example, to select an assessment device such as the *Portage Project Checklist* if your program stresses the cognitive skills delineated by Piaget, since few of these skills are included in the cognitive section of the *Portage Checklist.* To make such a choice means that the teacher will spend a lot of time measuring skills with little relevance to those taught.

Basically, six major curriculum or program models exist for the preschool handicapped population and most programs can be categorized into one of them (see Figure 2). In most cases, they are derived directly from the early intervention programs developed in the 1960's for disadvantaged preschoolers. All of these models differ in theoretical foundation, the amount of teacher-directedness, and the role of assessment. The descriptions which follow are brief and somewhat simplified. However, they give sufficient information for you to understand the general differences between the models.

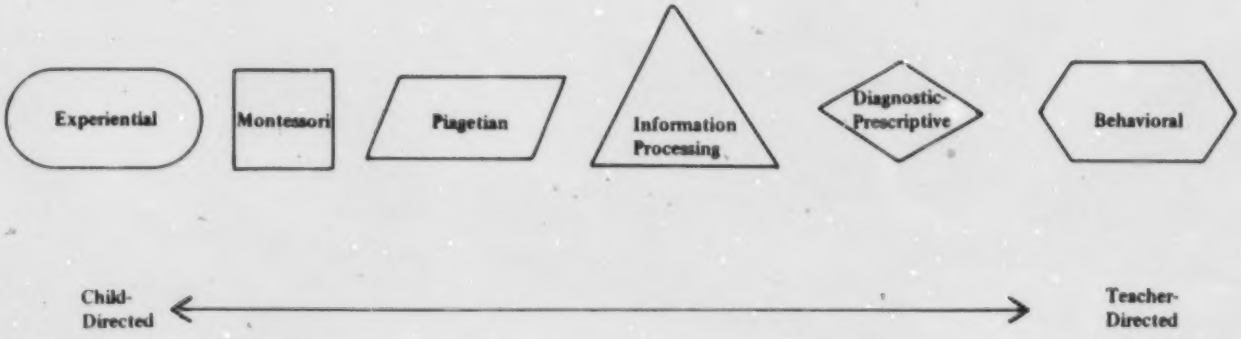
FIGURE 1

Comparison of Major Types of Assessment Techniques

	Purpose	Assumptions	Who administers	Norms/Reliability/Validity	Applicability to Intervention
NORM REFERENCED	1. To compare child's performance with performance of the children in the norm group who are usually the same age	1. Behavior generalizes across settings, sociocultures and observers 2. Single sampling of behavior is representative 3. Most measures use single score unidimensional 4. Maturation determined by genetic endowment 5. Similarity of previous experience	1. Psychologists - Psychometricians 2. Clinicians 3. Therapists 4. Diagnostic Specialist Administrator must be well trained	1. Usually dominant culture norms 2. Statistical norms 3. Must have high degree of validity and reliability a) Devices are stable and consistent measure of content b) Devices actually measure what they intend to measure	1. Teach children skills of dominant culture 2. Remediate processes 3. Usually only indicates general or global problems, thus often not helpful in planning daily intervention
CRITERION REFERENCED	1. To compare child's performance to a set of standards or criteria not to the performance of other children	1. Behavior is multi-dimensional 2. Behavior is situation specific 3. Repeated instances are being related or observed 4. Behavior is influenced by experience	1. Teachers 2. Sometimes paraprofessionals 3. Other professional personnel Administrator need not be specifically trained	1. No Norms - just a set of standards (sometimes taken from normative tests) 2. Reliability and validity linked to showing progress *within* an individual 3. Content validity crucial because it becomes the basis for child's program	1. Devices aid in determining the level of functioning and point out skills which the child needs to develop, thus facilitating program development through the delineation of skills present and absent 2. Content of devices focus on motor, cognitive, language, social-emotional and self-help skills
OBSERVATION	1. To watch the child in the natural environment without personally interacting with him and without making demands or structuring his response	1. The environment - people situation, etc. - influences the child's behavior 2. Behavior should be judged/analyzed within the content that it occurs 3. Behavior is complex and has many subtleties and must be observed over time or different situations	1. Teachers 2. Other professionals 3. Those collecting data as a part of a behavior modification program must be well trained	1. No norms - idiographic 2. Societal expectations 3. People who use the less structural techniques do not deal with validity and reliability 4. Behaviorists get interrater reliability and assume validity because actual behavior is observed	1. Applicable to intervention, not only delineates skills but also gives contextual information that helps in designing learning activities and sequences
PIAGETIAN	1. To investigate the effects of various encounters upon the rate of development which is not viewed as fixed 2. To ascertain the quality or level of response to the task in several dimensions	1. Behavior/development is the interaction of the child (who possesses certain competencies) with the environment 2. Competencies are interrelated and change as child interacts with environment 3. Competence is a hierarchical organization of abilities with several relatively independent branches 4. The quality and level of development at one stage depends heavily upon the level of skill acquisition at an earlier stage	1. These devices require skills in clinical observation and probing 2. The devices are constructed so that other professionals can administer but administrators must understand purposes and assumptions of devices	1. Use of Case Study methodology 2. Validity and reliability viewed in the traditional sense but with smaller Ns	1. Helps by determining the level of functioning and the quality of functioning 2. Environment is used to teach and develop skills 3. Content of devices focus on cognitive intervention

Figure 2

A Continuum of Early Childhood Curriculum Models



Experiential Model. This approach is based upon the maturational theory of child development and the educational philosophies of Dewey and Froebel. One of the assumptions of the model is that if stimulating experiences are provided for the child, he will grow and learn. In other words, the child is free to explore and structure his own learning through discovery which, it is hoped, will motivate him to learn and also will allow him to go at his own pace. Most of the preschool and nursery programs for normal children in this country use this approach. The preschool environment is arranged so that many educational materials and activities are available to the child.

Few programs for the handicapped use the experiential model, however, and very few materials are available. Most of the ones that are available were developed for disadvantaged four- and five-year olds. Some of the early programs for the hearing impaired, which stress language development in the context of everyday experiences, fit into this category. During routine and play activities, the mother provides the child with language. These experiences are supposed to stimulate language development, hence the term *language stimulation.*

Assessment in the experiential approach is rather informal. In programs for the handicapped, the child may receive some type of formal assessment with norm-referenced devices prior to being admitted into the program. Later, informal observation by the teacher, therapist, or parent is used to determine how the child is progressing and what kinds of experiences he or she needs.

Montessori Model. Maria Montessori believed that children learn through motorically manipulating objects. The structure of her programs, therefore, is found primarily in materials and routine activities which she developed for different levels of difficulty. Many of these items are now found in preschools all over the country and referred to as "educational materials." Montessori believed that many routine activities should have a structured motoric component in order to improve the child's skill and memory. For example, there is a proper structure and sequence to *wiping a table* with a wet cloth. The child is supposed to begin in the upper left-hand corner and move the cloth straight across to the right-hand corner. The child then returns the cloth to the left-hand corner and repeats this straight wiping motion until he or she reaches the lower right-hand side of the table.

The role of assessment in this model is an interesting proposition. If assessment procedures are to remain in harmony with the philosophical approach, teacher observation or informal assessment and checklists could measure a child's developmental level in terms of his ability to use particular materials. For example, there could be different developmental levels for the child's play with blocks, sand, paint, etc. The teacher could observe and record the level in relationship to various types of materials.

There are a number of Montessori preschool programs in this country; they serve children primarily from middle- and upper-income levels who are nonhandicapped. Two new programs are attempting to translate the principles of Montessori into programs for preschool handicapped children. They are in the early stages of developing assessment methodology. Perhaps, at some point they will have assessment and curriculum materials to share with others who want to use this approach. The following is a list of the two programs using the Montessori approach:

(1) Language and Speech Therapy Program
East River Montessori School
577 Grand Street
New York, NY 10002

(2) Side By Side
Franklin County Educational Collaborative
359 Main Street
Greenfield, MA 10301

Piagetian Model. This model is based upon Jean Piaget's theory of cognitive development. The focus is primarily upon assessing and teaching *cognitive* skills and processes. Fine and gross motor skills, receptive and expressive language, social skills, or self-help skills can not be measured with cognitively-oriented assessment devices. The cognitive skills delineated by Piaget are as follows: a) *sensorimotor stage*—visual pursuit and localization, object permanence, development of means for achieving environmental events, development of causality, spatial relationships, development of imitation; b) *preoperational stage*—classification, seriation, number, measurement, and space.

The first programs to use the Piagetian model contained mildly handicapped children. More recently, a number of talented professionals such as Diane Bricker, (University of Oregon), Beth Stephens, (University of Texas), and Cordelia Robinson (Meyer Children's Rehabilitation Institute) have developed procedures for using Piaget's theory with more severely handicapped children. It is possible that although these programs were developed for the severely handicapped, they could be useful with moderately handicapped preschoolers.

The following is a list of devices utilizing Piagetian theory.

1. *Assessment In Infancy: Ordinal Scales of Psychological Development*
I. Uzgiris and J. Hunt
University of Illinois Press
University of Illinois
Urbana, IL 61801
(Age 0-2 years)

2. *The Stages of Sensori-Motor Intelligence In the Child From Birth to Two Years*
I. Caseti and I. Lezine (UCLA Revision; Kopp and Sigman)
UCLA Department of Pediatrics & Psychiatry
Los Angeles, CA 90024

3. *Alber: Einstein Scales of Sensori-Motor Development*
Department of Psychiatry
Albert Einstein College of Medicine
New York, NY
(Age 0-2 years)

4. *Schaumburg: Early Childhood Developmental Milestones of Intellectual Development Checklist* (SEEC) [Sensory Motor and Preoperational]
Schaumburg Early Education Center
524 E. Schaumburg Road
Schaumburg, IL 60194
(Age 0-7 years)

These *programs* for handicapped preschoolers use the Piagetian approach.

5. High/Scope First Chance Outreach
Michael Hanes, Director
600 North River Street
Ypsilanti, MI 48197
• uses the *Cognitively Oriented Curriculum* (Weikart, et al.)

6. Cordelia Robinson
Meyer Children's Rehabilitation Institute
444 South 44th Street
Omaha, NB 68131
• developed an assessment device which is a modification of the *Uzgiris-Hunt*

7. Early Intervention with High Risk Children
Jeff Seibert
Mailman Center for Child Development
Debbie School
Post Office Box 520006
Miami, FL 33152
• developed an assessment device to be used in the classroom

This *article* describes assessment and intervention strategies for the severely handicapped using a Piagetian approach:

> Stephens, Beth. "A Piagetian Approach to Curriculum Development for the Severely, Profoundly and Multiply Handicapped." In E. Sontag (Ed.), *Educational Programming for the Severely and Profoundly Handicapped.* Reston, VA: Council for Exceptional Children, 1977.

Information Processing Model. Osgood's model of learning assumes that children bring in information [input] on visual, auditory, and tactile channels. The information is processed in the brain [integration]. Then the child responds through the verbal or motoric (tactile) channels [output]. The *Illinois Test of Psycholinguistic Ability* (ITPA) is the best known assessment device associated with this model. The subtests of the ITPA are as follows: auditory reception, visual reception, verbal expression, manual expression, auditory association, visual association, auditory sequential memory, visual sequential memory, auditory closure, grammatic closure, and visual closure.

The *Goal: Language Development Program 1* (Karnes, 1972), developed for use with low-income preschoolers, was one of the first curriculums to use the information processing model. It is based upon the subtests of the ITPA and contains short criterion-referenced checklists as part of the curriculum.

More recently, two prominent programs have used the model to develop assessment devices for preschool handicapped children: the Precise Early Education of Children with Handicaps (PEECH) project and Project First Step: Early Recognition and Intervention.

The following is a list of information processing assessment devices:

1. *Illinois Test of Psycholinguistic Ability* (for ages 2-10)
   S. Kirk, J. McCarthy, W. Kirk 1968
   University of Illinois Press
   University of Illinois
   Urbana, IL 61801

2. *Classroom Assessment: Systematic Child Observation Assessment Programming* (SCOAP) (Field Trial Edition)
   Assessment Manual: Systematic Child Observation Assessment Programming
   M. B. Karnes
   Colonel Wolfe School
   University of Illinois
   403 E. Healey
   Champaign, IL 61820

3. *Preschool Screening System* (Field Trial Edition) (ages 4.4 - 5.4)
   P. Hainsworth and M. Hainsworth 1974
   P. S. S.
   Box 1635
   Pawtucket, RI 02862

4. *Meeting Street School Screening Test* (ages 5.0 - 7.6)
   P. Hainsworth and E. M. Sigueland (1969)
   Meeting Street School
   667 Waterman Avenue
   East Providence, RI 02914

5. *Information Processing Curriculum Assessment Device*
   J. Seigel, P. Hainsworth, and M. Hainsworth
   P. S. S.
   Box 1635
   Pawtucket, RI 02862

Diagnostic-Prescriptive Model.* Currently, most of the programs for preschool handicapped children fall into this category which includes assessment as an integral part of the teaching-learning cycle. Each child receives a diagnostic assessment with norm-referenced devices prior to program placement. Once in the program, he receives an educational assessment which employs primarily a criterion-referenced assessment tool supplemented by information from observation. This assessment identifies the child's performance level and indicates those skills he needs to work on next. Using the information from the criterion-referenced assessment, the teacher writes individual behavioral objectives, and then selects and finally teaches the lesson to the child. Figure 3 demonstrates the process used in the diagnostic-prescriptive model. It is a continuous cycle. The teacher assesses each child's individual needs in as much depth and as specifically as possible. He or she then plays a major part in directing the learning activity for the child.

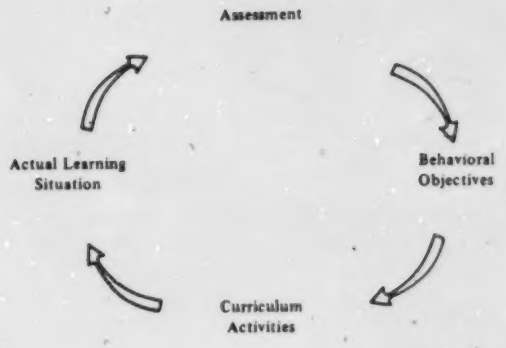


Figure 3
Diagnostic-Prescriptive Process

The content for the diagnostic-prescriptive model is based upon the developmental-maturational theory which maintains that normal development is sequential; hence, for different ages there are a set of normally expected developmental milestones. Children are assessed with a criterion-referenced checklist which contains developmental milestones in the following areas: gross and fine motor, social-emotional, self-help, receptive and expressive language, and cognitive skills.

The content of the criterion-referenced devices associated with this model is similar because the developers took items from the same instruments, i.e., the *Gessell, Stanford-Binet,* the *Bayley,* the *REEL,* etc. The weakest part of these devices is typically the cognitive section. It is made up of items from intelligence and other tests which are not purely cognitive. Therefore, items in this section actually measure vocabulary, fine motor skills, perceptual motor skills as well as a few cognitive skills. The devices do tend to vary in two ways: (1) format for presenting skills and (2) length or comprehensiveness.

The first of two common formats lists skills in each developmental area by age. For example, on the *Portage Project Checklist* all of the different kinds of

*Editor's note - While the term "diagnostic" is used with this model, it does not mean "diagnosis" as defined in the introduction to Section II and in Chapter 2. Actually it refers to both diagnosis and assessment as defined there.

motor skills are listed together—ages birth to one, one to two years, etc. Subskills are mixed together. For example, between three and four years are listed cutting, walking up and down stairs, throwing, pedaling a tricycle, swings, slides down slide, catches ball, etc. This provides a picture of all of the different motor skills (developmental milestones) a child should normally master between three and four years of age. Other criterion-referenced assessments which use this same format are: *Learning Accomplishment Profile* or LAP (Sanford, 1975); *Carolina Developmental Profile* (Lillie and Harbin, 1976); *Behavioral Developmental Profile* (Donahue, et al., 1975); *Lexington Developmental Scale* (Irwin and Coleman, 1975); and *Memphis Comprehensive Development Scales* (Quick, et al., 1974).

The second or other type of format is the task analysis of subskills within a large developmental skill area. For example, the *Manual for the Assessment of The Deaf-Blind Multiply Handicapped Child* (Collins and Rudolph, 1975) is divided into large developmental skill areas such as self-help, language, cognitive, social, and motor. Self-help is then divided into subskills like eating. Eating, itself, is divided into several component subskill areas such as food textures, bottle feeding, use of spoon, etc. In turn, each one of these is task-analyzed. This format pinpoints specific performance levels within a specific task. This type of format is particularly helpful to those working with moderately to severely handicapped children. The task analysis format is also helpful to teachers who have little training in the area of preschool handicapped education or little knowledge of child development. Other devices which use a similar format are: Behavioral Characteristics Progression or BCP (Santa Cruz County School, 1973); and Callier-Azusa Scales (Stillman, 1975).

The final way in which criterion-referenced assessment devices vary is in the length. As a result of more public schools beginning programs for preschool handicapped children and trying to serve as many children as possible with limited staff and resources, some programs have twelve children, one teacher and one aide in the morning and then the same teacher and aide in the afternoon *but twelve different children.* The teacher has a limited amount of time for assessment activities; thus, he or she uses one of the devices that is shorter and takes less time to administer. The *Carolina Developmental Profile* and the *Memphis Comprehensive Development Scales* are the two devices most used in this situation within the Diagnostic-Prescriptive Model. Both the *Carolina* and the *Memphis* have fewer items per age range than do most of the other devices. The format used by both is "skill listed by age." The longer devices using this format are the *LAP*; the *Behavioral Developmental Profile;* and the *Portage Project Checklist* (Bluma, et al., 1976).

Assessment devices using detailed task-analysis of subskills are longer and more specific. The following is a list of criterion-referenced checklists:

1. *BCP (Behavioral Characteristics Progression)*
   Santa Cruz County Schools, Santa Cruz, California
   VORT Corporation
   P. O. Box 11132
   Palo Alto, CA 94306

2. *Callier-Azusa Scales*
   R. Stillman (ed.) 1975
   Callier Center for Communications Disorders
   University of Texas at Dallas
   1966 Inwood Road
   Dallas, TX 75235

3. *Carolina Developmental Profile*
   D. Lillie and G. Harbin 1976
   Kaplan Press
   600 Jonestown Road
   Winston-Salem, NC 27103

4. *Marshalltown Behavioral Developmental Profile*
   Donahue, Montgomery, Keiser, Roecker, Smith 1975
   Marshalltown Project
   507 East Anson
   Marshalltown, IA 50158

5. *Developmental Therapy Objective Rating Form* (DTORF)
   M. Wood 1972
   Technical Assistance Office
   698 North Pope Street
   Athens, GA 30601

6. *Learning Accomplishment Profile*
   A. Sanford 1975
   Kaplan Press
   600 Jonestown Road
   Winston-Salem, NC 27103

7. *Lexington Developmental Scale*
   Irwin & Coleman, et al., 1974
   The Child Development Centers of U.C.P.B.
   P. O. Box 8003
   465 Springhill Drive
   Lexington, KY 40503

8. *Criterion Referenced Checklist*
   P. Tesauro & C. Takeshita
   Project Parent-Child
   Department of Speech Pathology and Audiology
   Speech and Hearing Center
   2450 South Vine
   University of Denver
   Denver, CO 80210

9. *A Manual for the Assessment of the Deaf-Blind Multiply Handicapped Child*
   M. Collins & T. Rudolph 1975
   Midwest Regional Center for Services to Deaf-

Blind Children
715 W. Willow Street
Lansing, MI 48906

10. *Portage Project Checklist*
Bluma, M. Shearer, Frohman, Hilliraid 1976
Portage Project
412 E. Slifer Street
Portage, WI 53901

11. *Memphis Comprehensive Development Scales*
Quick, Little, Campbell 1974
Fearon Publishers
6 Davis Drive
Belmont, CA 94002

12. *Pupil Progress Evaluation Plan*
J. Dickerson, M. Evanson, L. Spurlock
Panhandle Child Development Association, Inc.
421½ Sherman Avenue
Coeur d'Alene, ID 83814

13. *Comprehensive Early Education Profile*
King, Joyner, Smith, Thompson, Meredith, Bishop, Newquist, Brewer, Davis
Comprehensive Early Education Program
1720 7th Avenue South
Birmingham, AL 35294

Behavioral Model. This is perhaps the most structured and the most teacher-directed model. It is based primarily upon Skinner's (among others) theory o arning. The basic assumption is that all behavior is learned behavior; all children can learn if the right techniques are applied. This approach puts a great deal of emphasis on the *technology* of helping children learn and has been successful with the moderately to severely handicapped. It was also used successfully with disadvantaged preschoolers. Reinforcement, extinction, cueing, prompting, fading, shaping, and chaining are a few of the model's techniques. Many of the programs in the diagnostic-prescriptive category also use some of these behavioral techniques. They do not emphasize them, however, and are not as strict in their use as programs using the behavioral model—especially regarding the amount and specificity of data collected. With a few exceptions, most of the behavioral models were developed by professionals who were experts in the technology of behavior modification, but were not experts in early childhood curriculum content. Thus most of their assessment devices focus on skills that are obviously necessary and functional in nature. For example, many of the devices developed for the Behavioral Model have a strong self-help section. It does not take an expert in child development or curriculum development to recognize the importance of the development of eating, toileting, and dressing skills. On the other hand, there are important language, cognitive, and social skills which are critical but are not obvious to an untrained observer. Psycholinguists have taught us a great deal about the subtleties of language development—many of which are usually missing from the curriculum and assessment materials developed by behaviorists. In some cases, normal developmental milestones are used as a basis for the program and assessment device. In other cases, the assessment is imbedded within the curriculum.

Precise recording of behavior or the child's performance is imperative, and it is done by those trained in systematic data collection. Specificity and precision are two characteristics of the assessment procedures within the behavioral model.

The following programs are examples of this model.

1. Teaching Research Preschool
A Division of the Oregon State System of Higher Education
345 N. Monmouth
Monmouth, OR 97361

2. Fargo Preschool for the Handicapped
S.E. Mental Health and Retardation Center
700 First Avenue, South
Fargo, SD 58102

3. School for Contemporary Education
7201 Winstott Road
Springfield, VA 22151

In addition to devices for each of the curriculum models, people are usually interested in devices for three areas: language, social-emotional and perceptual skills. Some of these devices are discussed in the remainder of this chapter.

Language Assessment

There are basically three kinds of language assessment devices: 1) those that contain developmental milestones and language concepts; 2) psycholinguistic; and 3) information processing.

Language Concepts and Developmental Milestones. There are numerous assessment devices that measure language concepts and developmental milestones. All of the criterion-referenced devices listed under the diagnostic-prescriptive model fall into this category. These assessment devices assume that all children progress through a series of developmental milestones as they acquire language. Examples of items on this type of language assessment are: labels pictures; has vocabulary of 250-800 words; uses plurals; uses compound sentences; tells how simple objects are used; counts to three; etc. The following is a list of developmental-milestone language-assessment devices:

1. *The Bzoch-League Receptive-Expressive Emergent Language Scale* [REEL]
K. P. Bzoch and R. League
The Anhinga Press
420 Boulevard
P. O. Box 13501

Gainesville, FL 32604
Age: 0-3 years

2. *Birth - 3 Scale*
T. Bangs and S. Garrett
Speech and Hearing Institute
University of Texas
Health Science Center at Houston
1343 Meursund Avenue
Houston, TX 77035

3. *Houston Test for Language Development*
Houston Test Company
P. O. Box 35152
Houston, TX 77035
Age: Infancy - 6 years.

4. *Utah Test of Language Development*
M. Meacham, L. Jex, J. D. Jones
Communication Research Associates
Box 11012
Salt Lake City, UT 84111
Age: 9 months - 16 years

5. *Verbal Language Development Scale*
M. J. Meacham
American Guidance Service, Inc.
Publisher's Building
Circle Pines, MN 55014
Age: 1 month - 16 years

6. *VANE - L Scale*
J. Vane
Clinical Psychology Publishing Co., Inc.
Brandon, VT 05733
Age: 2½ - 6 years

Psycholinguistic. This type of assessment is based upon how children develop and use language structures such as prepositions, verbs, different sentence types, different types of questions, negatives, etc. This approach is the most complex, specific, and comprehensive in terms of delineating skills and their sequence. As a consequence, it is time-consuming and requires skilled and trained personnel to administer. The following are psycholinguistic devices.

1. *Development Sentence Analysis*
L. Lee
Northwestern University Press
1735 Benson Avenue
Evanston, IL 60201
Age: 2 words - complete sentences

2. *Northwestern Syntax Screening Test*
L. Lee
Northwestern University Press
1735 Benson Avenue
Evanston, IL 60201
Age: 3-8 years

3. *Carrow Elicited Language Inventory*
E. Carrow
Learning Concepts
2501 N. Lamar
Austin, TX 78705
Age: 3-8 years (Spanish and English)

4. *Test for Auditory Comprehension of Language*
E. Carrow
Learning Concepts
2501 N. Lamar
Austin, TX 78705
Age: 3-8 years (Spanish and English)

Information Processing. This approach is based upon the information processing model. The following device falls into this category. (See Information Processing Section.)

Illinois Test of Psycholinguistic Abilities
S. Kirk, J. McCarthy, W. Kirk
University of Illinois Press
University of Illinois
Urbana, IL 61801
Age: 2-10 years

Social-Emotional Assessment

These are probably the least well-developed of all assessment devices. We are still at a beginning stage in our ability both to understand and delineate specifically how children develop social and emotional skills. Although some theories about how young children develop emotionally exist, the assessment procedures at the present time, can only be used by trained clinicians and are primarily geared toward identification of pathology. All of the assessment devices listed for the diagnostic-prescriptive and behavioral models have sections with assess social-emotional skills. The following is a list of other devices which assess these skills.

1. *Preschool Behavior Questionnaire*
[A Screening Device]
L. Behar and S. Stringfield (1974)
Linc Press
1006 Lamond Avenue
Durham, NC 27701
* This device is to be used in day care, nursery school, and preschool settings in order to identify children with behavior problems.

2. *Developmental Therapy Objective Rating Form* (DTORF)
M. Wood 1972
Technical Assistance Office
698 North Pope Street
Athens, GA 30601
Age: 3-16 years
* This device was developed to be used with emotionally disturbed children.

3. *School for Contemporary Education Social Skills Checklist*
   7201 Winstott Road
   Springfield, VA 22151
   * This is a program using a behavioral approach with emotionally disturbed and behaviorally disordered children.

4. *California Preschool Social Competency Scale*
   Levine, Elzey, Lewis 1969
   Consulting Psychologist Press
   577 College Avenue
   Palo Alto, CA 94306
   Age: 2-5 years
   * This is a norm-referenced device which is especially useful in measuring skills necessary for success in a classroom.

5. *The Psychiatric Behavior Scale*
   W. F. Barker, L. Sandler, A. Borneman, G. Knight, F. Humphries, S. Rison
   Louise Sandler
   Center for Preschool Services
   The Franklin Institute Research Laboratories
   20th and Race Streets
   Philadelphia, PA 19103
   Age: 2-6½ years
   * This device assesses the preschool child's emotional development.

Perceptual Assessment

Increasingly, many professionals are realizing the importance of the development of perceptual or visual skills during the preschool years. Many theorists contend that perceptual skills and, in many instances, the lack of perceptual skills or distortions in the way a child perceives his or her world directly influence the ability to develop language, motor, cognitive and, in some instances, social-emotional skills. Professionals using the Piagetian and information-processing approaches are especially interested in the measurement of perceptual skills; however, each approach measures particular perceptual skills according to their theoretical approach. Sometimes these skills are the same and sometimes they are markedly different. The content of any perceptual assessment device you choose must be examined to determine if it is compatible with the philosophy of your program and measures what you want to measure.

Many professionals are already aware of two old and well-known devices: The *Bender Motor Gestalt Test* (L. Bender, 1938); and the *Frostig Developmental Test of Visual Perception* (M. Frostig, 1961). Over the last few years, both of these devices have received a great deal of criticism concerning their reliability and validity. Professionals have been cautioned to be extremely careful in the interpretation of results due to the lack of accuracy in the scores (Salvia and Yellesdyke, 1978).

Two other devices which assess perceptual skills are the *Carolina Developmental Profile* and *Psychoeducational Evaluation of the Preschool Child.* The *Carolina* is a criterion-referenced assessment device, one section of which measures perceptual skills. The *Psychoeducational Evaluation* is based upon the assessment techniques of Else Hoeussermann. It is similar to Piagetian assessment in that both the existence of skills and the quality of skill development are being examined. This device contains a section which assesses perceptual skills and other sections which assess cognitive, motor, and language development.

Psychoeducational Evaluation of the Preschool Child
E. Jedrysek
Grune and Stratton, Inc.
111 Fifth Avenue
New York, NY 10003

As we learn more, the devices to measure perceptual skills should become more numerous and more sophisticated.

Summary

It is hoped that listing and describing devices by categories will be helpful to those who are trying to examine, analyze and select an assessment instrument that will: 1) provide the kind of information needed on each child, 2) match the characteristics of the children, and 3) be consistent with the goals and philosophical approach of the program.

CHAPTER 4
Mildly to Moderately Handicapped Preschoolers: How do you select program evaluation instruments?
by Robert Sheehan

The selection of instruments for measuring the progress of mild to moderately handicapped preschool children is an important issue for early childhood special educators. Information from these instruments helps teachers plan appropriate and needed experiences for children and provides audiences such as the project staff, administrative and funding agencies, and parents with information regarding the effectiveness of the total program.

In order to select appropriate instruments, educators need to examine their programs carefully, determine what their information needs are, and then select the instruments that best provide that information. In this chapter, considerations for selecting appropriate instruments will be discussed.

The first distinction that needs to be made with regard to instruments for this population of children is in regard to their purpose. In early childhood programs for handicapped children, many types of instruments are used for quite different purposes, e.g., screening, diagnosis, assessment, evaluation, etc. Some devices have singular uses while others, notably assessment instruments, serve multiple purposes (e.g., planning instructional programs and evaluation).*

The second distinction that needs to be made with regard to instrumentation is that instruments do not need to be defined in any restrictive way. I prefer to define an *instrument*, not as a tool published by a testing corporation but, as *any attempt to gather information systematically*. An instrument has two basic characteristics. One, it should gather information specifically related to its purpose(s): e.g., measuring change in areas in which instruction was provided. Two, information yielded by one person must be qualitatively similar to the information produced by another person using the same instrument on the same subject. This definition allows formal tests, developmental checklists, rating scales, observation schedules, and other items to be considered instruments.

Much of the discussion in this book is directed toward formal tests and developmental checklists. This is because the field of early childhood special education has utilized these approaches most widely. However, there are also many other excellent procedures, which are viable for collecting data for evaluation purposes.

Principles

While I could suggest what I feel are the best instruments for you to use, this would be doing you and the children with whom you work a disservice. Early childhood special education personnel must make their own decisions about instruments based on the specific nature of their program and the children they are serving. Therefore, I can be of most help in suggesting principles to consider and discussing instruments that I have found useful in other early childhood special education settings.

The first principle is that *your choice of instruments should be directly related to the type of argument that you wish to make in your evaluation.* This argument should reflect the nature of your program, the handicapping condition of the children you are serving, as well as your intended audience. I will not dwell on this principle too much as it is discussed further in Chapter 11. Let me suggest, however, that before you talk to a measurement "expert," know what areas you want to measure, why, and for whom.

The second principle is that *your decision to use an instrument should include a consideration of the psychometric properties of that instrument.* You should be concerned with the reliability and validity of any measurement procedures you use. Unfortunately, very few instruments currently employed in early childhood special education have demonstrated reliability and validity for the population of children served. For example, the fact that an instrument was developed by drawing items from several standardized tests does not mean that it is reliable or valid or that it provides any more information than existing standardized procedures. If nonstandardized procedures are to be used, I suggest the use of one or two standardized instruments, which assess similar areas, for evaluation purposes.

The third principle is that *your choice of instrument should reflect your choice of evaluation design.* If the design necessitates an instrument which has norms, then obviously you cannot choose a non-normed instrument. If your evaluation design requires repeated data collection you must be sure that the instrument you select does not require too much time to administer and that repeated assessments do not cause children to react strangely.

*Editor's note–Chapter 3 deals primarily with instruments for assessment, while this chapter is devoted to instruments for evaluation.

The Instruments

You can find a discussion of instruments useful for early childhood special education in several sources. Among the most informative is a text by Cross and Goin (1977) entitled *Identifying Handicapped Children: A Guide to Casefinding, Screening, Diagnosis, Assessment, and Evaluation.* This text contains a bibliography which is an excellent summary of instruments in use in early childhood special education programs.

A second source of instruments is a series of publications released by the Center for the Study of Evaluation, UCLA Graduate School of Education (Hoepfner, Stern, and Nummedal, 1972). These publications present a thorough review of the psychometric properties of many instruments found in preschools and kindergartens across the country. Many of these instruments may be useful for early childhood special education programs.

A third source of instruments is the *Handicapped Children's Early Education Program Overview and Directory.* This document, which is produced each year, lists the instruments used for screening and assessment by all HCEEP demonstration projects. Table 1 provides a listing of the instruments more commonly used by the 127 demonstration projects funded during 1978-79.

Table 1

Name of Instrument	Number of Projects Using the Instrument
Bayley Scales of Infant Development	36
Learning Accomplishment Profile (LAP)	32
Denver Developmental Screening Test	23
Bzoch-League Receptive-Expressive Emergent Language Scale (REEL)	18
McCarthy Scales for Children	18
Peabody Picture Vocabulary Test (PPVT)	16
Portage Behavior Checklist	15
Vineland Social Maturity Scale	15
Alpern-Boll Developmental Profile	12
Standford Binet	11

There are several comments which I must make about the use of these and other devices for evaluation purposes. There are no tests in existence which were specifically designed for the purpose of evaluation. There are many tests that were designed for screening (such as the *Developmental Profile,* 1972) and for diagnosis (such as the *Illinois Test of Psycholinguistic Ability* or ITPA, 1968). There are also instruments that have been developed for program planning purposes (such as the *Learning Accomplishment Profile* or LAP, 1975). Unfortunately, early childhood educators cannot look in the manual of an instrument and decide whether the designer of that instrument has determined whether it is useful for evaluation purposes. The educator must make that decision alone.

There are, however, special problems that I believe should be considered regarding two commonly used types of devices. One problem in selecting a *screening* instrument for evaluation purposes is that the instrument may not be responsive to changes which occur in children during the year. By definition, a screening instrument is administered in a short period of time, which means there are relatively few items. Consequently, there are also fewer areas in which to show change over the period of time during which the intervention is occurring.

A second type of problem is encountered when programs use an *intelligence test* such as the *Stanford-Binet* (1972) for evalution purposes. These tests were designed to assess the psychological construct of intelligence. As such, they have been specifically developed to be resistant to differences in performance attributable to minor changes in environment. In other words, they assess the stable aspects of growth and development and yield a very conservative estimate of a program's overall impact on handicapped children's development.

In my activities with early childhood special education programs, I rarely advise programs to select a single instrument for evaluation purposes. What I do is examine the stated goals of the program, the nature of the children's handicapping conditions, and the resources of the program for collecting data; then I advise the development of an instrument battery that best addresses all of these variables. Frequently, this battery includes one or two instruments which reflect language development, one or two which reflect cognitive growth, a behavior rating instrument, a visual perception instrument, and at times, a project-developed instrument.

There are in existence several instrument packages which tap all or most of these areas. These include the *Comprehensive Identification Process* or CIP (1975) and CIRCUS (1974). As far as individual instruments or less comprehensive packages are concerned, I feel very positive about the *McCarthy Scales of Children's Abilities* (1972), the *Stanford-Binet* (1972), the *Metropolitan Readiness Test* (1976) and the *Sequenced Inventory of Communication Development* (1975), to name a few. While none of these devices have accurate psychometric data from populations of children similar to those served in early childhood special education programs, they are ones that I have used and that I feel are constructed adequately, and that reflect many goals traditionally found in such programs. Citations for these and other instruments are listed below.

List of Instruments

Developmental Profile (G. Alpern and T. Boll). Indianapolis: Psychological Development Publications, 1972.

Illinois Test of Psycholinguistic Ability (S. Kirk, J. McCarthy, and W. Kirk). Urbana: The University of Illinois Press, 1968.

Learning Accomplishment Profile (A. Sanford). Winston-Salem, N. C.: Kaplan School Supply, 1975.

Stanford-Binet Intelligence Test (L. M. Terman, M. A. Merrill, and R. L. Thorndike). Atlanta: Houghton Mifflin, 1972.

Comprehensive Identification Process (R. R. Zehrbach). Bensenville: Scholastic Testing Service, 1975.

Circus (S. Anderson, G. Bogatz, T. Draper, A. Junglebut, G. Sedwell, W. Ward, and A. Yates). Princeton: ETS, 1974.

McCarthy Scales of Children's Abilities (D. McCarthy). New York: The Psychological Corporation, 1972.

Metropolitan Readiness Test (G. Hildreth, N. Griffiths, and M. McGauvran). Atlanta: The Psychological Corporation, 1976.

Sequenced Inventory of Communication Development (D. Hedrick, E. Prather, A. Tobin). Seattle: University of Washington, 1975.

Summary

We are at a time in our profession when we are being asked to do more than we can adequately document. With the support of such endeavors as the Bureau of Education for the Handicapped's four Early Childhood Research Institutes, we will hopefully witness the development and testing of appropriate instruments for this purpose. Several other organizations and agencies are also taking on this task. Ultimately, young handicapped children will be served and will benefit from such assistance. As Sheldon White (Goldman and Lewis, 1976) has said: "We mount our programs on behalf of children using ordinary human judgment and intuition. We take guidance from data when there is such guidance. We get along without formal data for that large number of questions for which formal data do not provide answers" (pp. 38-39). I would only add to White's comments by suggesting that we not be too quick to abandon our search for appropriate instruments, but rather accept it as a challenge in which all early childhood special educators can engage.

References

Cross, L., and Goin, K. *Identifying Handicapped Children: A Guide to Casefinding, Screening, Diagnosis, Assessment, and Evaluation.* New York: Walker, 1977.

Goldman, K. S., and Lewis, M. *Child Care and Public Policy: A Case Study.* Princeton: Educational Testing Service, 1976.

Hoepfner, R., Stern, C., and Nummedal, S. G. *CSE ECRC Preschool/Kindergarten Test Evaluations.* Los Angeles, CA: UCLA Graduate School of Education, 1972.

Technical Assistance Development System (TADS) and Western States Technical Assistance Resource (WESTAR). *Handicapped Children's Early Education Program: 1978-79 Overview and Directory.* Chapel Hill, NC, and Seattle: TADS and WESTAR, 1979.

CHAPTER 5

Mildly to Moderately Handicapped Infants: What should influence your approach to measurement?

by Claire B. Kopp

Almost every topic related to early intervention is surrounded by questions, issues, and even sharp disagreements. The subject of measurement, which has proven to be particularly challenging, is no exception. Ostensibly, the problems center on how change is to be documented, but even more fundamental issues regarding theoretical and methodological questions arise. Solutions to these dilemmas are not easily found although numerous approaches have been proposed. Often these stem from an intervener's preference for specific growth models, view of the child's teachability, and philosophical perspectives.

My own philosophical perspective, governed by a developmental orientation, is outlined in the first portion of this paper. This is followed by a two-part discussion of some general ideas about measurement. In the first part, the reader will find the focus on intervener-parent communication with a measurement orientation derived from a counseling or public health model. There is virtually no emphasis on measuring progress in the young infant. It is argued that this is the method of choice when infants are very young and/or when parents are new to intervention programs. The second part is devoted to measurement *vis a vis* the older infant or young child. The paper concludes with discussion of additional techniques of measurement with specific reference to the mild to moderately handicapped young child.

A Philosophical Perspective

Developmental Issues. From a conceptual viewpoint I find it useful to employ a developmental orientation as a starting point to understand child development and the issues related to assessment and intervention. Such a perspective argues for a dynamic, transformational, multivariate view of the child. For example, infancy and early childhood are characterized by extensive changes in the behavioral repertoire with new behaviors acquired, others refined and reorganized, and still others disappearing (e.g., babbling). Moreover, it is not just that change occurs, but rather that the numerous capabilities of the young child undergo qualitative as well as quantitative shifts (Flavell, 1972; Piaget, 1952, 1954, 1970; Werner, 1948; Wohlwill, 1973).

Clearly, a developmental perspective may lead to interest in the transformational process in and of itself (McCall, 1979; Wohlwill, 1973). From a more applied stance, consideration of transformations allows one to pinpoint different behavioral clusters associated with specific ages (e.g., competence at age two versus competence at age four). Moreover, the developmental perspective helps us systematize our thoughts about organization and sequential acquisition of abilities, refinement of skills, periods of qualitative change, and so forth. Measurement of the young handicapped child should reflect an awareness of many of these developmental qualities.

My views on the growth, change, and teachability of children are based on developmental theory, psychiatry, psychobiology, ethology, and evolutionary perspectives (see, for example, Brunner, 1972; Erikson, 1950; Flavell, 1972, 1977; Hinde, 1974; Piaget, 1952, 1954, 1970; Scarr-Salapatek, 1976), as well as from research findings regarding normal and atypical infants and young children. For example, I have been impressed with the striking behavioral similarities reported in the literature among all kinds of infants—whether they be normal or mild to moderately handicapped, reared in California or in the African bush (see reviews, Kopp, 1979; Kopp and Parmelee, 1979; Werner, 1972). Similarly, I have been influenced by findings which demonstrate the extraordinarily nonplastic nature of early developmental sequences (Cicchetti and Sroufe, 1976; Dennis, 1941; Fantz, Fagan and Miranda, 1975). At the cross-species level of research, I am intrigued by data that suggest that cats and monkeys show object permanence (Gruber, Girgus and Banuazizi, 1971; Wise, Wise and Zimmerman, 1974) and that primates and older infants demonstrate many similar primitive problem-solving strategies (Kohler, 1925; Piaget, 1952, 1954). In like manner, I have been guided by other data that depict infants as far *less* flexible, adaptive, and capable of learning than older children. Limitations have been found in levels of learning (Hinde, 1974), problem solving (Case, 1978), memory (Brown, 1975), and self-differentiation (Mahler, Pine, and Bergman, 1975).

Biological Aspects. All of these data support schools of thought which argue that biological factors are the major influencers of very early development. These factors are the basis for behavioral sequences and change (Brunner, 1972; Flavell, 1972; Piaget, 1970; Scarr-Salapatek, 1976; Wohlwill, 1973). Scarr-Salapatek, for example, argues that evolutionary and biologically-determined pressures and mechanisms predispose infants to develop many behaviors typical of their species. These behaviors have adaptive and survival value and thus need to be acquired and maintained whatever the rearing conditions or stress to the organism.

The environment, she and others suggest, neither *causes* behaviors to appear nor modifies their essential form; the environment simply provides "materials and opportunities" for learning and development. However, she notes that evolutionary pressures led to human rearing environments that have many common elements (objects, people, events), so that almost all types of early rearing experiences will, sooner or later, lead to attainment of the sensorimotor repertoire. Since infants cannot control the rearing environments they are exposed to, it is fortunate indeed that, during a period of relative helplessness, many kinds of settings lead to the same end product.

The Environment. Given a genetic blueprint that guides the form and nature of early development, it is also apparent that at some point, not precisely delineated, biological influences attenuate and environmental influences show increasing and more specific effects. Thus, the child becomes more directly modifiable as a function of characteristics of the milieu. While some suggest this transition starts toward the latter part of the first year, others speculate it is most likely to be found toward the end of the sensorimotor stage and the beginning of representational and analytic thought. Irrespective of the exact period, research findings indicate that in the second year, disparities in child capabilities become evident as a function of training (McGraw, 1975), social class (Golden, Birns, and Bridger, 1973), and maternal style (Clarke-Stewart, 1973).

The Caregiver. The brief resume presented above has focused exclusively on the nature of the infant and young child. This, however, cannot stand apart from consideration of the role of the environment and more specifically, that of caregivers. In brief, recognizing the force of biological mediators but also aware of the environment's role as an impetus to growth, it seems that the caregiver's first and continuing role is to provide a sensitive "caring" background. This involves understanding the infant and meeting physical and psychological needs. In essence, caregiving should help infants maintain a positive motivational set toward their environments.

A positive affective set is considered by White (1959) to be part of "effectance motivation" which is defined as the desire *to do* and to be effective. Although White considers efficacy to be intrinsic to the organism, he also states it can be influenced by supportive interactions. These, for example, may involve caregiver responsiveness and provision of opportunities for the infant to have new experiences (Goldberg, 1977). Clearly, White and Goldberg's points of view are reminiscent of Erikson's (1950) eloquent theme of "mutal trust" being the *sine qua non* of infant caregiver interactions.

Although empiric validation of these constructs are not currently available, I find the substance of the ideas have great intuitive appeal. Surely, feeling comfortable and secure in one's surroundings is an important impetus for learning and exploration. What better foundation could infants have?

Parenthetically, Goldberg (1977) extended the notion of efficacy to parent behavior, proposing that caregivers' feelings of their own efficacy derive from evaluations of interactions with their infants. Evaluations, in part, stem from perceptions of specific infant characteristics and behavior, including infant readability (clearly defined signals), predictability (reliability of behavior occurrence), and responsiveness (quality and extent of infant reactions).

Returning to discussion of the environment: a second and equally important role of the parent involves *engineering an appropriate milieu* that sets the stage for channeling and expanding existing capabilities of the child. Here, reference is made to physical and social environments that, for example, promote use of language, foster play, develop awareness of social rules, and the like. Similarly, it is probably beneficial to have caregivers devise situations that constrain numerous diffuse, nondirected explorations of the child, and help him/her focus on salient features of the milieu. Based on the arguments presented earlier, it is likely that caregiver planning and channeling become increasingly important during the child's second and third years, and beyond. Few data exist that are pertinent to this particular idea; nonetheless, research yields incontrovertible evidence that different environments have profound and lasting effects on the older child's capability.

Thus far, I have elaborated a point of view which visualizes biological mechanisms controlling the form and function of early life. These are presumed to operate with environmental mediators which foster development. I have suggested that one vital way of promoting early infant development is via caregiver provision of a supportive milieu. This should facilitate exploration (at whatever level the infant is capable of) and concomitantly maintain a positive affective set. [I might add, as an aside, that fostering development is sometimes viewed in the context of accelerating change processes, i.e., the precocious infant. This approach has little appeal to me, since longitudinal data repeatedly demonstrate little stability between infant developmental level and later cognitive skills (Bayley, 1933; McCall, Hogarty, and Hurlburt, 1972; McCall, Eichorn, and Hogarty, 1977).]

Measurement Issues

Parents. If caregivers are viewed as key figures first in helping infants establish trusting relationships and positive motivational sets, and second, in channeling and promoting child and development, then in the face of handicapping conditions, interventions and of course measurements initially should be directed primarily toward parents and secondarily toward the infant. The idea of drawing parents into intervention programs is hardly novel. Indeed, the clinical literature and empirical evaluations suggest that programs for children at sociocultural risk *which included parent components* demonstrated more lasting effects than programs without such content. However, what I am suggesting goes beyond parent participation: we need a major intervention commitment to parents of handicapped children. In this plan,

parents are guided to center stage, while infants are temporarily assigned to the cast of supporting players.

Why is a commitment to parents necessary? First, because the realization that a child is handicapped leads to feelings of doubt, uncertainty, and anxiety (Roskies, 1972). Often, these feelings last for a long time. Therefore, one focus of intervention needs to be communication—an open forum in which parents feel free to ask questions, voice doubts as well as certainties, and ventilate feelings.

On a second level, commitment is necessary in order to provide parents with numerous opportunities to learn about development of children in general and specifically about children with handicapping conditions. Given time, parents realize that in many ways handicapped and normal infants and young children are similar in needs and capabilities.

Finally, a commitment to parents is important to help them understand their own child. Here, the focus should be on establishing and maintaining a mutually trusting relationship. Earlier, reference was made to the feelings of efficacy of young children and parents (Goldberg, 1977; White, 1959). What interferes with this motivational set? The clinical literature suggests that infants who are handicapped, developmentally delayed, or motorically involved send signals that are unclear and poorly defined (Jones, 1977; Roskies, 1972). Although many parents in time learn to interpret signals correctly, still others have more difficulty and misinterpret infant responses. Not surprisingly, confusion, nonresponsiveness, and feelings of inadequacy result (Emde and Brown, 1978). But, by talking and focusing on child characteristics, parents can begin to resolve these problem areas.

In sum, the type of initial intervention approach that is outlined above is focused on processes involved in promoting sensitive parenting. Its goal is to establish a foundation for subsequent parent-child interactions. Presumably, in time each parent will feel free to establish a style that "feels right" for him or her. Recognize that the perspective outlined above neither demands parents to be extensions of interveners nor asks them to focus initially on specific sequences of activities that will lead to specific infant outcomes. Also recognize that many similar ideas have already been put into operation by Fraiberg (1971) and her team of associates as they worked with parents of blind infants.

Even with this type of parent intervention, it is essential to derive a measure of program effectiveness. An appropriate model might be based on counseling or public health perspectives in which program effectiveness is broadly defined by the clients' use of available services. Thus, with parents of the handicapped, documentation could be made by quantifying parents' voluntary use of intervention services, willingness to ask questions, number of appointments kept, requests for assistance in making/buying toys or modifying the home environment for the child, and parent willingness to help other parents—in short, any set of indices that informs the intervener that communication is either open and ongoing or not occurring. If the latter situation should arise, the intervener must be prepared to try additional ways to reach parents.

I personally feel less comfortable with measurement approaches that might be aimed at evaluating "good caregiving." Far too few data exist, and far too many caregiving styles have been described, to attempt to quantify something as idiosyncratic and qualitative as "good." Moreover, a public health approach puts the onus of accountability on the intervener where it rightly belongs. It demands that interveners commit themselves to an interaction with parents. This is not necessarily the case when a primary measurement focus centers on an evaluation of "good" caregiving. Indirectly, parents may become the source of accountability and receive the blame for program inadequacies.

After parents and interveners develop a communication system, and after parents and infants establish a trusting, efficacious mode of interaction, then it is appropriate to change the focus of intervention and documentation. At this point, interveners should direct attention explicitly and primarily to the characteristics of the handicapped child with a goal of fostering additional growth. Implementation of an intervention program and concurrent documentation of child progress become critical issues. Some of these issues are explored in the next section.

Child. If one views the nature of the young child from a developmental perspective, then it follows that measurement of progress and, of course, intervention should be approached similarly. I recognize that this view poses formidable, albeit ultimately solvable, problems.

As noted earlier, measurement techniques need to reflect the transformational, multivariate, hierarchical, and organizational characteristics of early development. However, current techniques for the most part mirror only a limited portion of these characteristics. A few of the more common documentation approaches are discussed in the following paragraphs. Then, some modifications are suggested that may move us closer to a more inclusive perspective than has heretofore existed.

At present, two prominent approaches to measurement can be identified. One involves evaluation of *performance of discrete behaviors* of the child: e.g., discriminates circle from square, holds crayon, etc. Often, this type of measurement system coincides with use of a curriculum by which the intervener has attempted to teach an infant or young child a particular skill or ability. Thus, the ability is considered to be the referent or criterion goal.

One advantage of this approach is that it allows interveners to tally the child's accomplishments quickly. Alternatively, a major disadvantage is that interveners tend to view the behavior as an isolated function and thereby ignore possible relationships with other skills. Stated another way, analyzing performance of a single discrete behavior often does not lead to evaluation of either the organization or the interactive nature of the child's repertoire. Thus, the intervener may not grasp the essential meaning and subsequent implications of the

child's success or failure on the referent task.

Another common approach to documentation involves *norm-referenced instruments* (Bayley, Griffiths, etc.) which are used to compute developmental quotients at levels. When two or more such measures are taken, (early into the program and six months later), the item of interest is essentially the *rate of change* demonstrated by the child. The advantage of this approach is that the intervener can compare a mild to moderately handicapped child's progress against a set of developmental expectations generated from the performance of a substantial number of normal children.

However, this approach has several pitfalls. First, it unduly "penalizes" some groups of handicapped children. For example, visual impairment imposes extraordinary challenges for young children in their attempts to learn about and organize their world. Given a supportive environment, blind infants do acquire knowledge, but frequently it takes them a longer period of time than sighted infants. Therefore, developmental delays may be noted, often signifying a different time frame for change rather than impairment *per se*. Documentation of progress assessed with normative instruments does *not* make this important distinction.

A second problem with the approach relates to what is included and omitted in the content of the tests. I suggest that many important behaviors are excluded, while some relatively "unimportant" behaviors are included and given undue emphasis. This, of course, will affect documentation of progress. Consider the fact that the normative approach has always argued for inclusion of a behavior within a set of standards on the basis of visibility; i.e., most children show particular kinds of behaviors at certain ages. Of course, this line of reasoning does have some face validity, since a majority of children do demonstrate similar age-related behaviors.

Is it important to demonstrate that handicapped children show adequate performance on all of these behaviors? I have an intuitive feeling that the probable answer is no. The heart of the problem lies in the fact that virtually no consideration has been given to the actual meaning of items in the tests. To my knowledge, there has been no analysis of the developmental functions and processes that are tapped by one or another item. I think that any analysis of this kind would demonstrate that developmental-normative tests tap a very limited domain of functions and processes of the older infant's repertoire.

Extending the argument further, it is clear that most normative scales were generated long before the emergence of recent and innovative thinking about early development. Not surprisingly, the scales have limited inclusion of categories that we now consider important, such as child play and exploration, selective and sustained attention, problem-solving strategies, recognition and recall memory, use of others as resources, imitation, and a host of other functions.

Some of these topics have a conceptual basis in Piagetian theory (1952, 1954, 1970), whereas others stem from attempts to define and describe the multiple abilities associated with developmental competencies (Bronson, 1971, 1974; White and Watts, 1973). These new conceptualizations and associated research have had significant implications for furthering our understanding of normal young children. This knowledge should be applied to the young handicapped child and, of course, should enter the domain of measurement.

Finally, it seems that documentation approaches do not take into account the type of developmental change that occurs in infancy and childhood. This issue bears on questions of quantitative and qualitative processes of change. In turn, it relates to intervention goals and whether one is attempting *to expand* a certain type of ability (arithmetic skills) or alternatively *is fostering* acquisition of new behaviors (object permanence).

Some processes of change are more prevalent during one age period than another. For example, during infancy there is considerable evidence of acquisition—smile, prehension, stranger wariness, object permanence, verbal language, and the like. Simultaneously, there is only a limited amount of expansion of skills. In contrast, during the school years childrens' abilities are characterized by expansion—number of facts acquired, amount of vocabulary learned, and so forth—and not as much by acquisition.

Consideration of these processes of change (Flavell, 1972), as well as the nature of the organism should help us determine indices that are particularly appropriate for measurement. The possibilities are numerous; however, they do require awareness of the nature of the child.

These general comments about measurement are a prelude to the final paragraphs of this paper. The following are ideas about specific indices that could be used to enhance current documentation techniques with handicapped two- to three-year-old children.

Additional Measurement Techniques

While one could choose any number of indices to add to existing measurement instruments, choices stem from two reference points: contemporary literature on child competencies (Block and Block, 1973, 1978; Bronson, 1971, 1974; Mahler, Pine, and Bergman, 1975; Anderson and Messick, 1974; White and Watts, 1973) and the clinical literature which describes problem areas of handicapped and "at risk" children (Birch and Gussow, 1970; Chess, 1974, 1976; Kessler, 1970).

Dealing with People and Objects. Around the age of two years, children should begin to develop strategies or ploys to deal with objects and people. This involves planning and organizing a sequence of activities, carrying them to fruition, and evaluating the consequences. Naturally, at age two years, the manifestations of these behaviors will be at a very primitive level. One way to examine problem-solving strategies with objects is via use of sensorimotor series. For example, the Casati-Lezine (Kopp and Sigman, 1972) sensorimotor series includes a problem-solving task that can elicit multiple strategies because it is cognitively demanding (Kopp, O'Connor, and Finger, 1975). It would be relatively simple to document changes

in the child's use of strategies, the conditions that foster growth of strategies, and factors that elicit disorganized approaches.

A similar tactic can be employed to evaluate the young child's developing strategy skills with peers and caregivers. Here, reference is made to the child's ability to make appropriate overtures to adults for assistance or to peers for sharing and playing. White and Watts (1973) offer some interesting clues about techniques that could be used for measurement.

Play. Play is another area that deserves to be explored as an additional documentation technique. Fortunately a rich and extensive conceptual literature exists on early play (Brunner, 1973; Brunner, Jolly and Silva, 1976; Garvey, 1977; Piaget, 1962) which is complemented by a series of empirical studies (Fenson, Kagan, Kearsley, Zelazo, 1976; Kagan, Zelazo and Kearsley, 1977). The findings yield data on the form of developmental progression of play and also attest to individual differences. Thus, categorization of play could produce simple checklists or more comprehensive analyses of type of play demonstrated, number of schemes employed, length of time a dominant form of play was used, and the like.

The one important point to recognize concerns the issue of individual differences in play. Questions have been raised about the implications of limited versus extended amounts of play demonstrated by any given child. A prudent course suggests avoidance of extreme statements about the significance of amount of play.

Attending. Lastly, I suggest interveners would find it useful to measure a child's ability to maintain attention to a task or play activity. The clinical literature suggests this is a problem with young handicapped children. Moreover, data suggest problems with attention are a prominent feature of learning disabilities (Ross, 1976).

Unfortunately, empirical data are sparse with respect to attention and young children; indeed, the age period between nine and thirty months is virtually ignored. Notwithstanding, perceptive clinicians and interveners can identify children who flitter (visually and/or motorically) from activity to activity. For example, I recently observed a two-year-old in an intervention program who visually connected to a play activity for about forty-five seconds. After each short bout of attending, he averted his eyes, dropped his head, and started to move his arms and legs. A meaningful documentation approach for this child would include measures of "time on" and "time off" for selected activities. This, of course, should coincide with intervention attempts to expand the duration of sustained attention. A simple kind of graph could be used to record progress and illustrate change.

In sum, I have suggested some ideas that might be translated into documentation techniques for the young handicapped child. Clearly, others might propose equally valid ideas and alternatives. Numerous alternatives should be explored although one limiting factor should always be considered. That factor concerns the underlying conceptual rationale for the alternative. If the rationale is weak, it is likely that a technique based upon the rationale will also be weak. Although we need to expand our repertoire of techniques that are used to document progress, we should proceed with caution.

References

Anderson, S., and Messick, S. "Social Competency in Young Children." *Developmental Psychology* 10 (1974): 282-293.

Bayley, N. "Mental Growth During the First Three Years: A Developmental Study of Sixty-one Children by Repeated Tests." *Genetic Psychology Monographs* 14 (1933): 1-92.

Birch, H. G., and Gussow, J. D. *Disadvantaged Children: Health, Nutrition, and Social Failure*. New York: Harcourt, Brace and World, 1970.

Block, J., and Block, J. H. "A Longitudinal Study of Personality and Cognitive Development in Young Children." *A Summary Report Prepared For Participating Families*, 1978.

Block, J., and Block, J. H. "Ego Development and the Provenance of Thought, A Longitudinal Study of Ego and Cognitive Development in Young Children." *Progress Report for NIMH Grant Number MH 16080*. Washington, DC: National Institute of Mental Health, 1973.

Bronson, W. C. "Competence and the Growth of Personality." In K. S. Connally and J. S. Bruner, (Eds.), *The Growth of Competence*. New York: Academic Press, 1974.

Bronson, W. C. "The Growth of Competence: Issues of Conceptualization and Measurement." In H. R. Schaffer (Ed.), *The Origins of Human Social Relation*. New York: Academic Press, 1971.

Brown, A. L. "The Development of Memory: Knowing, Knowing about Knowing, Knowing How to Know." In H. W. Reese (Ed.), *Advances in Child Development and Behavior* (Vol. 10). New York: Academic Press, 1975.

Bruner, J. S. "The Nature and Uses of Immaturity." *American Psychologist* 27 (1972): 1-22.

Bruner, J. S. "Organization of Early Skilled Action." *Child Development* 44 (1973): 1-11.

Bruner, J. S. Jolly, A., and Sylva, K. *Play. Its Role in Development and Evolution*. New York: Bases, 1976.

Case, R. "Intellectual Development from Birth to Adulthood: A Neo-Piagetian Interpretation." In R. S. Siegler (Ed.), *Children's Thinking: What Develops?* Hillsdale: Erlbaum, 1978.

Chess, S. "Factors Affecting the Development and Evaluation of Handicapped Children." In E. Schopler and R. J. Reichler, *Psychopathology and Child Development: Research and Treatment.* New York: Plenum, 1976.

Chess, S. "The Influence of Defect on Development in Children with Congenital Rubella." *Merrill-Palmer Quarterly of Behavior and Development* 20 (1974): 255-274.

Cicchetti, D., and Sroufe, L. A. "The Emotional Development of the Infant with Down's Syndrome." *Child Development* 47 (1976): 920-929.

Clarke-Stewart, A. "Interactions Between Mothers and Their Young Children: Characteristics and Consequences." *Monographs of the Society of Research in Child Development* 38 (1973) (6-7, Serial No. 153).

Dennis, W. "Infant Development under Conditions of Restricted Practice and of Minimum Social Stimulation." *Genetic Psychology Monographs* 23 (1941): 143-189.

Emde, R. N., and Brown, C. "Adaptation After the Birth of a Down's Syndrome Infant: A Study of Six Cases, Illustrating Differences in Development and the Countermovement Between Grieving and Maternal Attachment." *Journal of The American Academy of Child Psychiatry* 17 (1978): 299-323.

Erikson, E. H. *Childhood and Society.* New York: W. W. Norton & Company, 1950.

Fantz, R. L., Fagan, J. F., and Miranda, S. B. "Early Visual Selectivity." In L. Cohen and P. Salapatek (Eds.), *Infant Perception: From Sensation to Cognition, Vol. I. Basic Visual Processes.* New York: Academic Press, 1975.

Fenson, L. Kagan, J., Kearsley, R. B., and Zelazo, P. R. "The Developmental Progression of Manipulative Play in the First Two Years." *Child Development* 47 (1976): 232-236.

Flavell, J. H. "An Analysis of Cognitive-Developmental Sequences." *Genetic Psychology Monographs* 86 (1972): 279-350.

Fraiberg, S. "Intervention in Infancy: A Proposal for Blind Infants," *Journal of American Academy of Child Psychiatry* 10 (1971): 381-405.

Garvey, C. *Play.* Cambridge: Harvard University Press, 1977.

Goldberg, S. "Social Competence in Infancy: A Model of Parent-Infant Interaction." *Merrill-Palmer Quarterly* 28 (1977): 163-177.

Golden, M., Burns, B., and Bridger, W. "Review and Overview: Social Class and Cognitive Development." Presented at the Society for Research in Child Development, April 1973.

Gruber, H. E., Girgus, J. S., and Banuazizi, A. "The Development of Object Permanence in the Cat." *Developmental Psychology* 4 (1971): 9-15.

Hinde, R. A. *Biologic Bases of Human Social Behavior.* New York: McGraw-Hill, 1974.

Jones, O. H. M. Mother-Child Communications with Pre-Linguistic Down's Syndrome and Normal Infants." In H. R. Schaffer (Ed.), *Studies in Mother-Infant Interaction.* London: Academic Press, 1977.

Kagan, J., Kearsley, R. B., and Zelazo, P. R. *Infancy: Its Place in Human Development.* Cambridge, MA: Harvard University Press, 1978.

Kessler, J. W. "Contributions of the Mentally Retarded Toward a Theory of Cognitive Development." In J. Hellmuth (Ed.), *Cognitive Studies,* Vol. 1. New York: Brunner/Mazel, 1970.

Kohler, W. *The Mentality of Apes.* New York: Harcourt Brace, 1925.

Kopp, C. B. "Perspectives on Infant Motor System Development." In M. Bornstein and W. Kessen (Ed.), *Psychological Development from Infancy.* New York: Erlbaum, 1979.

Kopp, C. B., O'Connor, M., and Finger, I. "Task Characteristics and a Stage 6 Sensorimotor Problem." *Child Development* 46 (1975): 569-573.

Kopp, C. B., and Parmelee, A. H. "Pre and Perinatal Influences on Development." In J. Osofsky (Ed.), *Handbook of Infancy.* New York: Wiley, 1979.

Kopp, C. B., and Sigman, M. *The Stages of Sensori-Motor Intelligence in the Child from Birth to Two Years.* Casati, I., and Lezine, I. Paris: Center of Applied Psychology, 1968, (UCLA revision, 1972).

McCall, R. "Qualitative Transitions in Behavioral Development in the First Three Years of Life." In M. Bornstein and W. Kessen (Eds.), *Psychological Development from Infancy.* New York: Erlbaum, 1979.

McCall, B., Eichorn, H., and Hogarty, S. Transitions in Early Mental Development. *Monographs of the Society for Research in Child Development* 42 (1977) (Serial no. 171).

McCall, R. B., Hogarty, P. S., and Hurlburt, N. "Transitions in Infant Sensorimotor Development: The Prediction of Childhood I.Q." *American Psychologist* 27 (1972): 728-748.

McGraw, M. *Growth: A Study of Johnny and Jimmy.* New York: Appleton Century, 1975.

Mahler, M. S., Pine, F., and Bergman, A. *The Psychological Birth of the Human Infant.* New York: Basic Books, 1975.

Piaget, J. *Play, Dreams and Imitation.* New York: Norton, 1962.

Piaget, J. "Piaget's Theory," In P. H. Mussen (Ed.), *Carmichael's Manual of Child Psychology.* New York: Wiley and Sons, 1970.

Piaget, J. *The Origins of Intelligence in Children.* New York: International Universities Press, 1952.

Piaget, J. *The Construction of Reality in the Child.* New York: Basic Books, 1954.

Roskies, E. *Abnormality and Normality. The Mothering of Thalidomide Children.* Ithaca: Cornell University Press, 1972.

Ross, A. O. *Psychological Effects of Learning Disabilities.* New York: McGraw-Hill, 1976.

Scarr-Salapatek, S. "An Evolutionary Perspective on Infant Intelligence: Species Patterns and Individual Variations." In M. Lewis (Ed.), *Origins of Intelligence.* New York: Plenum Press, 1976.

Werner, E. "Infants Around the World: Cross-cultural Studies of Psychomotor Development from Birth to Two Years." *Journal of Cross-Cultural Psychology* 3 (1972): 84-112.

Werner, H. *Comparative Psychology of Mental Development* (Rev. Ed.). Chicago: Follett, 1948.

White, B. A., and Watts, J. C. *Experience and Environment: Major Influences on the Development of the Young Child.* Vol. 1. Englewood Cliffs: Prentice Hall, 1973.

White, R. W. Motivation Reconsidered: The Concept of Competence. *Psychological Review* 66 (1959): 297-333.

Wise, K. L., Wise, L. A., and Zimmerman, R. R. *Developmental Psychology* 10 (1974): 429-37.

Wohlwill, J. T. *The Study of Behavioral Development.* New York: Academic Press, 1973.

Zelazo, P., and Kearsley, R. "Functional Play: Evidence for a Cognitive Metamorphosis in the Year Old Infant." Paper presented at the Biennial Meeting of the Society for Research in Child Development, New Orleans, March 16-20, 1977.

CHAPTER 6

Mildly to Moderately Handicapped Infants: How do you assess the progress of both mother and infant?

by Earladeen Badger

Introduction

In discussing assessment procedures to measure treatment effects for mildly and moderately handicapped infants during their first twelve months of life, I recognize that I am treading on thin ice. There has been some well-deserved criticism (Lewis and McGurk, 1972; Elkind, 1973) of infant stimulation programs which attempt to measure the effects of specific intervention procedures with infant intelligence scales. Typically, there is a mismatch between the particular set of skills which the program seeks to emphasize and the instrumentation it used to test those skills. Further, there is no necessary continuity between intelligence as defined at one stage of development and as defined at another. Bayley (1970) has concluded: "The findings of early studies of mental growth of infants have been repeated sufficiently often so that it is now well established that test scores earned in the first year or two have relatively little predictive validity" (p. 1174). Lewis and McGurk's (1972) data as well as other information (King and Siegmiller, 1971) indicate that, even with respect to sensorimotor functions, there is a lack of continuity.

Given the caution regarding the predictive validity of infant assessment scales as well as the reports (Starr, 1971; Haskins et al., 1978) of the evanescent effects of infant intervention programs when treatment was terminated, it is perhaps not surprising that early identification and treatment of high-risk infants and their families are not widely accepted.

Over the past ten years, researchers have focused on two high-risk groups: infants who are at risk for environmental or cultural-familial retardation and infants with high potential for neuro-sensory damage. Premature or low birthweight infants make up a large proportion of this latter group. In analyzing early intervention studies, Meier, (1975) notes that a preventive approach is rarely taken. Less than three percent of the developmentally suspect population is referred for treatment under twelve months of age.

Coming from a program which is housed in a medical center and that has offered follow-up educational intervention for children during the first twelve months of age over the past six years, several problems of early identification and treatment have become apparent to us. (1) Except in cases of Down's Syndrome and severe mental retardation, which are apparent shortly after birth, there are a far greater number of handicapping conditions which are less obvious during the first year of life. (2) Physicians are still hesitant and parents often deny rather obvious developmental delays during the first year of life. (3) Aberrations of parenting as noted in immature, mentally retarded, or disturbed adolescents and/or adults are resistant to change and often produce infants who "fail to thrive" or who are psychologically and physically battered, and become "special" children at an early age.

Program Design

The population of infants in which we've intervened over the last six years includes both those who are at risk for environmental retardation (infants of inner-city, adolescent welfare mothers, sixteen years of age and younger) and those who are biologically vulnerable (premature, low birthweight infants). All infants are identified for inclusion in our follow-up educational intervention during the mother's lying-in period following delivery. Intervention consists of twenty postnatal mother-infant classes and weekly, biweekly, or monthly supportive home visits during the infant's first twelve months of life. A multidisciplinary staff includes an early childhood educator, a nurse, a social worker, paraprofessional aides, a neonatologist and physician's assistant with an occupational therapist, physical therapist, and psychiatrist available for consultation.

It is important to state that our program philosophy holds that the parent is the primary interventionist, which means that we believe that whether an infant thrives or not is determined in large part by the caretaking environment that the mother provides. Because of the recent acknowledgment and appreciation of the newborn infant's remarkable and appealing capacity for social interaction (Lozoff, et al., 1977; Brazelton, et al., 1975) our primary intervention strategy is to facilitate attachment behaviors. Our presence in the high-risk nursery allows us to soften the harsh effects of some hospital routines and accommodate the mother and infant's mutual needs for one another. Modern perinatal medical management, unfortunately, has replaced frequent nursing and abundant physical interaction with intermittent artificial feeding and minimal maternal contacts. It further imposes a pattern of early separation on mother and neonate. It seem appropriate that a program called "Infant Stimulation/Mother Training" (IS/MT) should attempt to ameliorate some of these adverse effects. The thrust of our intervention then is a developmental/behavioral approach. Mother is encouraged

to establish early emotional contact with her infant. As much as possible, she is involved in the physical care of her infant during his/her period of hospitalization. Upon her infant's discharge from the hospital, they are involved in weekly mother-infant classes. The logical and sequential steps in child development are presented in a class format and mother's behavior is "shaped" within the group as part of the group process and at home as the program is individualized for her infant. For those interested in a more detailed description of the class program, it is reported elsewhere (Badger, 1977).

Recognizing that the success of any intervention program must be assessed according to specific criteria related to the content of the program, we saw the need to assess the progress of both mother and infant. What I have chosen to share with you then is three instruments we've used to assess *maternal caretaking* ability and three assessments that we've used to measure *infant development.* Both of these sets of measures are important in that they complement each other in determining infant outcome. It is this interaction of constitutional and environmental variables which together over time are believed to influence development (Badger, Burns, and Vietze, 1978).

Maternal Caretaking Assessments

Maternal Risk Score (MRS). This instrument was designed by project staff to index the social environment of poverty mothers. As can be seen in Table 1, the MRS consists of fifteen items grouped into four categories: maturity of the mother, degree of planning, family support system, and level of poverty. Each item is scaled so that the highest weights reflect greatest risk. These items were selected in part on the observations of staff working with families and partly on an analysis of childrearing and family life patterns associated with poverty. This information is objectively quantifiable and easily retrievable through extant hospital records. In a paper which was presented at the Society for Pediatric Research meetings in New York in April, 1978, we reported on a sample of thirty-six mothers of first born infants. Their risk scores were computed at four intervals shortly after birth and at the time of their child's first, second, and third birthdays. This allowed for adjustments in a mother's risk score with her changing life situation and also permitted evaluation of the predictive value of the risk score at different points in her child's first three years of life. In this initial validation of a maternal caretaking assessment, maternal scores were found to have adequate stability across a three-year period and some predictive value for tests of their child's cognitive competence as measured by the *Bayley Infant Scales*—during the first two years of life. While the test of replication must occur in order to provide more evidence on the soundness of our approach, it is not difficult to see how an MRS instrument might be used in the absence of more detailed information to deploy services on a preventive basis rather than on the more prevalent crisis or remedial basis. This is especially true among those poverty families who are hard pressed to provide development-fostering experiences for a high risk infant.

Maternal Infant Interaction. This instrument was designed by project staff to assess changes in maternal caretaking ability during the first eight weeks of our postnatal mother-infant class prgram. As can be seen on this checklist instrument (Table 2), there are observations and recordings on five categories: mother's behavior in class, infant state, mother's behavior directed to baby, baby's behavior directed to mother, and nutritional component. In an article (Badger, Burns, and Rhoads, 1976) appearing in the *American Journal of Public Health*, we reported on a sample of thirty-nine mothers who had attended at least five of the first eight classes. A paired t-test analysis of the interaction scale data yielded a statistically significant difference between first- and eighth-week ratings on ten items. For example, mother participation as evidenced by attentiveness and relaxed behavior in class improved significantly ($p < .001$). Maternal behavior as evidenced by increased responsiveness towards infants also improved ($p < .001$) as did infant responsiveness to mother ($p < .001$). Although some of the interaction effect can be attributed to infant maturation, there was no significant change observed in infant state from class one to class eight. It may be inferred, therefore, that the mothers did, in fact, become increasingly responsive to the social cues provided by the babies. It should be noted that infants were all four to six weeks of age at the beginning of the observations and twelve to fourteen weeks of age when the observations were concluded. Inter-rater reliability in recording was 83 percent.

Home Visitor's Report. This instrument was developed by project staff as an adaptation of the Home Inventory (Elardo, Bradley, and Caldwell, 1975) in order to assess changes in maternal caretaking ability over a period of several months. As can be seen (Table 3), our home visitors record their weekly observations by attending to twenty-four items in four categories: infant state, infant responsiveness, home environment, and mother. Home visitors are instructed to check only those items actually observed during the visit.

In order to evaluate possible changes in maternal behavior as observed by the home visitor, a sample of eighteen mothers was randomly picked from a group of sixty-five mothers who had been recruited for participation in the service program.

Mother-infant pairs were distributed equally on the following bases:

(1) Mother-infant pairs who had completed the twenty-week class program and were also visited weekly in the home.
(2) Mother-infant pairs who never attended classes or who had low class attendance—less than five classes—but who were home visited on a weekly basis.
(3) Mothers who had two or more children in the home.
(4) Mothers who had infants weighing less than three pounds at birth.

TABLE 1

Maternal Risk Score

SCORE _____
MOTHER _____

Maturity of the Mother

1. ____ Mother's Age
   - 5 - 13 years and younger
   - 4 - 14
   - 3 - 15
   - 2 - 16
   - 1 - 17
   - 0 - 18 years and older

2. ____ Mother's Education (Last grade completed)
   - 4 - below 9th
   - 3 - 9th grade
   - 2 - 10th
   - 1 - 11th
   - 0 - 12th (H.S. diploma or G.E.D.)

3. ____ Educational commitment
   - 2 - not returning
   - 1 - Uncertain
   - 0 - Finished H.S. or has plans to return

4. ____ Prior type of schooling
   - 2 - Special (MR, Guilford, Dyer)
   - 0 - Regular

5. ____ Socially Deviant Behavior
   - 4 - ID as potential child abuser/C.P.S. Placement
   - 3 - criminal record
   - 2 - previous psychiatric history
   - 0 - no record

Degree of Planning

6. ____ Mother Has Other Children
   - 4 - 3 others under 5 years of age
   - 3 - 2 others under 5 years of age
   - 2 - 1 other (less than 2 yrs. spacing)
   - 1 - 1 other (2 or more years spacing)
   - 0 - No other children/or children over 5 years

7. ____ Prenatal Care for Mother
   - 3 - None
   - 2 - 1-2 visits
   - 1 - 3-5 visits
   - 0 - 6 or more visits

8. ____ Previous Abortion/Miscarriage
   - 1 - yes
   - 0 - no

Family Support System

9. ____ Caretaker
   - 3 - no plan/unlicensed day care
   - 2 - relative or friend
   - 1 - maternal grandmother
   - 0 - mother primary caretaker

10. ____ Living Arrangement
   - 3 - alone or with a friend
   - 2 - father or mother, or guardian, or relative
   - 1 - maternal grandmother
   - 0 - father of child, husband, or mother and father

Level of Poverty

11. ____ Geographic Area
   - 5 - West End, NW Fairmount, Over-the-Rhine, S. Cumminsville
   - 4 - Price Hill, Mt. Auburn, Winston Place, Avondale, Walnut Hill
   - 3 - East End, Evanston, Eastwood Village, Corryville
   - 2 - Northside, Fairview, Madison
   - 1 - Clifton, Bond Hill
   - 0 - Any other area

12. ____ Welfare Status of Mother
   - 3 - total family dependency on welfare
   - 2 - mother is totally dependent on AFCD
   - 1 - mother work/in training program - also gets AFDC
   - 0 - self supporting (no welfare)

13. ____ Telephone
   - 2 - no
   - 0 - yes

14. ____ Number of Mother's Younger Siblings in Household
   - 3 - over 6
   - 2 - 3-6
   - 1 - 1-2
   - 0 - none

15. ____ Does Mother have siblings with children 5 years in the home?
   - 2 - yes
   - 0 - no

Developed by E. Badger, Director, Infant Stimulation/Mother Training Project

47

TABLE 2

Mother-Infant Interaction Scale-Checklist

M = Most of the time 50%
S = Some of the time 25-50%
H = Hardly ever 25%

A. Mother's Behavior in class

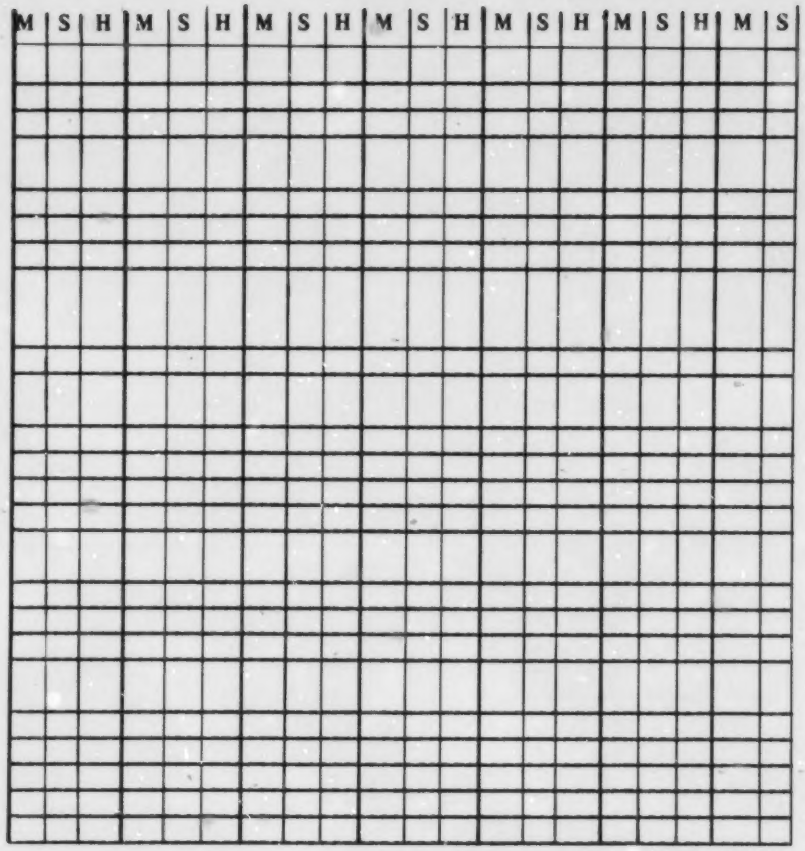
Attentive
Defensive
Relaxed

B. Infant State
Alert/Inactive
Alert/Active
Fussy/Crying

C. Mother's Behavior Directed to Baby
Attentive to Baby's Cues
Inappropriate Response/ behavior
Talks to Baby
Plays with Baby
Smiles at Baby
Looks at/watches Baby

D. Baby's Behavior to Mother
Looks at Mother
Vocalizes to Mother
Smiles at Mother

E. Nutritional Component
Mother Offers Bottle
Baby spits up
Mother offers pacifier
Mother offers Food
Mother offers juice/water



Infant Stimulation/Mother Training Project, U. C. College of Medicine, Cincinnati, Ohio 45267

TABLE 3

Infant Stimulation/Mother Training
Home Visitor's Report

Infant's Name: _____ Home Visitor: _____

Birthdate: _____ Date: _____

INFANT STATE: (Check one)

____ Alert/Active

____ Alert/Inactive

____ Sleepy/Difficult to Rouse

____ Fussy/Crying

INFANT RESPONSIVENESS:

1. ____ Responds to visual stimuli
2. ____ Responds to auditory stimuli
3. ____ Responds to tactile stimuli
4. ____ Coordinated body movement

HOME ENVIRONMENT:

5. ____ Appears safe and free of hazard
6. ____ Evidence of surrogate mothering from other family members
7. ____ Infant has jump seat, walker, high chair, or stroller
8. ____ Developmental toys available for infant

MOTHER:

9. ____ Talks to infant
10. ____ Responds to infant's vocalization with vocal or verbal response
11. ____ Initiates verbal interchanges with visitor–asks questions, spontaneous comments
12. ____ Expresses ideas easily and freely
13. ____ Praises infant's behavior or qualities
14. ____ Conveys positive feelings when talking about infant
15. ____ Caresses or kisses infant
16. ____ Shows emotional response to praise of infant by home visitor
17. ____ Provides verbal feedback on material covered in class
18. ____ Interacts with infant, using toy received in class
19. ____ Demonstrates muscle activity demonstrated by visitor
20. ____ Asks clarifying questions about material covered in class
21. ____ Repeats stimulation activity demonstrated by visitor
22. ____ Voices concern about infant's *slow* development
23. ____ Shares personal problems
24. ____ Complains about family problems

[Questions 9-17 are taken from HOME INTENTORY (Birth to Three) from The Center for Early Development and Education, 814 Sherman, Little Rock, Arkansas, 72202]

E. Badger, Director, Infant Stimulation/Mother Training
U. C. College of Medicine, Cincinnati, Ohio 45267

(5) Mothers who had infants weighing more than three pounds but less than five pounds at birth.

The average length of time families were visited at the time of this evaluation was five months. The mean number of visits completed was seventeen.

Since the primary goals and objectives of our educational program were measured in the Mother category, I will report briefly on changes in that area. There was a dramatic increase in the number of positive interactions observed by the home visitor during the initial month as compared to the fifth month. The initial mean was 17.3 items checked or approximately four items a week as opposed to a mean of 30.8 items during the fifth month, which is almost twice as many per week. While it can be argued that as infants mature they become more responsive, it is noted that infant state (active and inactive alert) was fairly constant from the first to the fifth months of home visitation. We therefore inferred that the mother's responsiveness to the teaching model provided by the home visitor positively influenced her interactions with her infant.

Infant Development Assessment

The two purposes for assessing infant development that we were mindful of are (1) to generate measures for program evaluation and (2) to identify prescriptive strategies for individual infants within a program. The variety of assessment approaches we've used has varied according to the populations served as well as the individual differences between handicapped infants. According to Simeonsson's analyses of the approaches, assessments can be looked at in terms of their foci—achievement, function, or competence (Simeonsson, 1977).

Achievement tests are downward extensions of the traditional intelligence scales for older children and adults. Infant development is assessed on the basis of passing or failing an achievement task. Examples of these scales are *The Bayley Scales of Infant Development, The Griffith's Mental Development Scale, The Denver Developmental Screening Test,* and *The Vineland Social Maturity Scale.* The advantages of achievement-based scales are that they are usually standardized, widely used, and have considerable acceptance. The limitations of these scales are that the standardization procedure usually excludes handicapped and at-risk infants and the manner in which they are administered is often incompatible with the handicaps unique to certain infants.

Functional assessments are increasing in use. They work well with behavioral objectives which can be developed from the IEPs or prescriptive programming for handicapped infants. The focus of the assessment is on the relationship between inputs (teacher stimuli) and outputs (infant response). For example, if you have baseline data on the infant's vocalization or socialization skills, you can make precise observations by using one of four different recording systems: frequency, duration, interval, or sampling. Impact of intervention can be easily demonstrated with this kind of assessment and the assessment can be tailor-made for an individual infant. The limitations of this assessment approach are that it is difficult to establish reliability of observation (training is required) and, since it is not intrinsically developmental in approach, it is difficult to examine preceding behavior or to predict subsequent behaviors.

Competency tests are an outgrowth of Piagetian theory and are finding increased application as an assessment technique. In this approach, the process rather than the product of cognitive function is assessed. The cognitive competencies assessed follow an invariant sequence in development. They are ordinal in nature but are not age-specified. In other words, the rate of development of infants may vary, but the sequence remains constant. The particular strategies an infant uses in imitative and intentional behaviors, as well as those he uses in dealing with space, time, object permanence and causality, are all examples of competencies which can be differentiated according to stage of development. Representative examples of these scales are *The Uzgiris-Hunt Infant Ordinal Scales of Psychological Development* and the *Albert Einstein Scales of Sensory-Motor Intelligence.* The advantages of this assessment approach are that it focuses on the individual infant, and it provides a firm basis for identifying competencies achieved and those yet to be acquired. Limitations which must be considered are that there must be a familiarity with and commitment to a theoretical perspective and the actual administration demands time and effort because optimal rather than elicited performance, as with a standardized test, is being assessed.

I'd like now to share briefly with you how we've used three different infant assessments in our educational follow-up program for high-risk infants during their first twelve months of life.

Uzgiris-Hunt Infant Scales. We administered the *Uzgiris-Hunt Scales* to document the risk status of infants born to adolescent mothers (sixteen years and younger) and to validate the treatment effect of our postnatal weekly mother-infant class program. In comparing six- and twelve-month mean composite testing scores of nine infants of young mothers attending weekly classes with a matched group of infants of young mothers in a home-visited comparison group, an analysis of variance indicated a significant difference in infant performance favoring the class group of infants at both six and twelve months of age. These data are described in greater detail elsewhere (Badger, Elsass, and Sutherland, 1974).

It is important to mention that all mothers were present while the tester administered the scales on a monthly basis during their infants' first twelve months of life. Also, class and home-visited infants received the same stimulation toys and materials as incentive for their participation in the program. However, mothers of class infants received specific instructions in play activities with their infants which would foster the development of the cognitive competencies assessed. I devised a series of sequential mother-infant games (Badger, 1971) for each of the six scales and further adapted the "Development of Schemas

Scale" in the *Infant and Toddler Learning Program* (Badger, 1971). We did not choose to assess our infant stimulation program with the *Uzgiris-Hunt Scales* without providing specific curriculum inputs. These two curricula provided direction to the mothers in arranging learning experiences which would enhance their infants' cognitive development. Since the stimulation activities are sequential in nature and not age-specified, they can easily be translated into behavioral objectives and IEPs for individual infants.

Bayley Infant Scales. We decided to administer the *Bayley Scales* to the same population of infants of adolescent mothers when the infants were twelve months of age. This is a norm-referenced test which has a high level of communicability as it produces an IQ score. The infants of young mothers in the home-visited group had a mean mental score of 79 which was significantly less ($p < .05$) than the mean score of 99 obtained by infants of young mothers in the class group. The mean motor scores of class infants was likewise significantly better ($p < .05$) than those in the home-visited group. These data are reported elsewhere (Badger, Elsass, and Sutherland, 1974).

Both the *Uzgiris-Hunt Scales* and the *Bayley Scales* assess intellectual or cognitive development. One is a criterion-referenced test and the other is a norm-referenced test. Our experience suggests that there is a high level of agreement between the two instruments. Both instruments take approximately an hour to administer. The *Uzgiris-Hunt Scales* can be mastered and administered by teaching staff as part of the intervention program. The *Bayley Scales* require a high level of training to administer reliably, and you would do well to engage a certified tester if you want to report on your results.

EMI Assessment Scale. We have recently begun to use the Field Test Version of the *EMI* (Elder, 1975) i.e., "education for multihandicapped infants," in our BEH follow-up program for premature infants. We have selected this instrument because, like the *Learning Accomplishment Profile* or *LAP*, it is easily adapted by the educator for individualized diagnostic and prescriptive teaching of infants. We selected it in preference to the *LAP* because it is a little less cumbersome for paraprofessional staff to implement. The scale provides for evaluation in five areas of development: gross motor skills, fine motor skills, socialization, cognition, and language. It consists of three behavioral items in each area for each month from birth through twenty-four months. The selection, sequencing, and normative assignment of items are based on information from the *Bayley, Uzgiris-Hunt, Gesell, DDST,* and a half dozen other developmental assessment instruments.

We do not have any data to report yet on our use of the EMI, but staff and parents are comfortable using it as they begin to develop IEPs for prescriptive teaching of premature infants. Since this population of infants is typically delayed in both motor and social development, we felt the EMI would provide a more comprehensive assessment of especially high-risk infants.

Summary

In assessing treatment effect or intervention outcomes with mildly and moderately handicapped infants, the instrument selected or developed should reflect the goals of your intervention efforts. The rationale behind the development of IEPs and behavioral objectives is to provide a framework for educational intervention which has as its foundation skill-fostering experience in normal child development. Depending on the handicapping condition of individual infants, the educational interventionist and parent decide together what treatment plan is to be followed. Periodic assessments of defined behavioral goals may be more meaningful than administering a standardized instrument in a vacuum. If parents are included as primary interventionists with their infants, and you have a treatment plan to operationalize their involvement, you would do well to assess the quality of their input.

In sharing the assessment instruments which we have used in the IS/MT Program in Cincinnati, you will note that we measure parental adequacy in addition to infant development. We believe that they are intimately entwined during the infant's first twelve months of life, and that maternal caretaking ability may prove to be the critical variable in determining infant outcome.

References

Badger, E. "The Infant Stimulation/Mother Training Project." In B. Caldwell and D. Stedman (Eds.), *Infant Education: A Guide for Helping Handicapped Children in the First Three Years.* New York: Walker, 1977.

Badger, E. *Mother-Infant Game.* ERIC Clearinghouse for Early Childhood Education. Urbana: University of Illinois, 1971 (Ed. 046510).

Badger, E. *Teaching Guides: Infant and Toddler Learning Program.* Paoli, Pennsylvania: The Instructo Corporation, 1971.

Badger, E., Elsass, S., and Sutherland, J. "Mother Training as a Means of Accelerating Childhood Development in a High-Risk Population." Paper presented at Society for Pediatric Research, Washington, DC, May 2, 1974.

Badger, E., Burns, D., and Rhoads, B. "Education for Adolescent Mothers in a Hospital Setting." *American Journal of Public Health* 66 (1976): 469-472.

Badger, E., Burns, D., and Vietze, P. "Maternal Risk Factors as Predictors of Developmental Outcome in Early Childhood." Paper Presented at Society of Pediatric Research, New York, April 17, 1978.

Bayley, N. *Carmichael's Manual of Child Psychology*, P. H. Mussen (Ed.). New York: Wiley, 1970.

Brazelton, T. B., et al. "Early Mother-Infant Reciprocity." In *Parent-Infant Interaction*, Ciba Foundation Symposium 33. Amsterdam: ASP, 1975.

Elardo, R., Bradley, R., and Caldwell, B. "The Relation of Infants' Home Environment to Mental Test Performance From Six to Thirty-Six Months: A Longitudinal Analysis." *Child Development* 46 (1975): 71.

Elder, W. "The EMI Assessment Scale." Charlottesville, VA: University of Virginia Medical Center, (Box 232), 1975.

Elkind, D. "Infant Intelligence." *American Journal of the Disadvantaged Child* 126 (1973): 143-144.

Haskins, R., Finkelstein, N., and Stedman, D. *Pediatric Annals* 7 (1978): 99-128.

King, W., and Siegmiller, B. "Cognitive Development from 14-22 Months of Age in Black, Male, First Born Infants Assessed by the Bayley and Uzgiris-Hunt Scales." Paper presented at a meeting of the Society for Research in Child Development, Minneapolis, 1971.

Lewis, M., and McGurk, H. "Evaluation of Infant Intelligence." *Science* 178 (1972): 1174-1177.

Lozoff, B., et al. "The Mother-Newborn Relationship: Limits of Adaptability." *The Journal of Pediatrics* 91 (1977): 1, 1-12.

Meier, J. "Screening, Assessment, and Intervention for the Young Child at Developmental Risk." In B. Z. Friedlander, G. M. Sterritt, and G. E. Kirk (Eds.), *Exceptional Infant* (Volume 3), New York: Brunner/Mazel, 1975 (pp. 605-651).

Sameroff, A. J., and Chandler, M. J. "Reproductive Risk and the Continuum of Caretaking Casualty." In F. D. Horowitz, M. Hetherington, S. Scarr-Salapatek, and G. Siegel (Eds.), *Review of Child Development Research* (Volume 4). Chicago: University of Chicago Press, 1975 (pp. 187-244).

Simeonsson, R. "Infant Assessment and Developmental Handicaps." In B. Caldwell and D. Stedman (Eds.), *Infant Education: A Guide for Helping Handicapped Children in the First Three Years.* New York: Walker, 1977.

Starr, R. "Cognitive Development in Infancy: Assessment, Acceleration and Actualization." *Merrill-Pamer Quarterly of Behavior and Development* 17 (1971): 153-186.

CHAPTER 7

Severely and Profoundly Handicapped: How do you measure sensorimotor and cognitive skills?

by Cordelia C. Robinson

The focus of this chapter is the assessment of sensorimotor and cognitive development in the young child who is severely or profoundly retarded. The most important criteria in assessing the cognitive development of such children is the selection of tasks or situations that will yield information that is educationally prescriptive. This bias on my part comes from several years of experience in attempting to use norm-referenced developmental assessments as a basis for educational planning and description of developmental progress with infants who are severely or profoundly retarded. The information derived from such assessments generally was of little value and added frustration to what was already a very difficult situation for child, parent, and evaluator. In this chapter, I will describe the rationale for an alternative basis for assessment using the *Uzgiris-Hunt Ordinal Scales of Psychological Development* (1975).

Population

The group about whom I am writing is severely to profoundly retarded young children who are between birth and seven years of age, and whose developmental or intelligence quotients are in the range of five to six standard deviations below the mean, or below 40. Within this group, we are likely to find a number of children who have multiple handicaps, either sensory or physical in nature, in addition to their retardation. This multiplicity of handicaps frequently makes it very difficult to determine the extent to which the delayed learning observed in the children is contributed to or caused by their physical or sensory disabilities. They are children who do not receive the typical range of sensory and motor experiences enjoyed by the average child in interaction with his social and physical environment. They are children whose opportunities for learning that "they can make interesting things happen" are likely to be limited; thus, they are likely to experience a cycle of diminished *motivation* for learning coupled with diminished *opportunities* for learning basic sensorimotor skills. These characteristics of the group of children known as severely and profoundly retarded seriously limit the utility of norm-referenced developmental scales as a basis for:

(1) predicting future developmental level;
(2) designing prescriptive educational environments; or
(3) evaluating the success of educational interventions with these children.

Rationale for the Developmental Model

The movement toward early educational and therapeutic intervention with children who are handicapped has relied heavily on a developmental model as the basis for planning and evaluating interventions. Resources such as Friedlander, Sterritt, and Kirk (1975) and Tjossem (1976) provide ample documentation of this point. Basically, the developmental model has involved use of information from child development literature in several domains (motor, cognitive, language and social development) as "maps" to be followed in the design of curriculum sequences for children with handicaps. Frequently, the details for those curriculum sequences have been derived from norm-referenced intelligence tests such as the *Bayley Scales of Infant Development* (1969) or developmental checklists such as those developed by the Portage and Marshalltown projects. These sequences have generally been determined in one of two ways. Some, such as the *Bayley Scales,* have been determined empirically by administering a pool of items to a number of children and arranging acceptable items in sequential order based upon the average age at which the item is passed. Others have been developed by "borrowing items" from a number of different sources and then ordering them on the basis of the data sources from which they were borrowed. A more extensive discussion of the implications of this approach with respect to the purposes of evaluation can be found in Chapter 11. Despite some measurement problems, the developmental model has served educators well in developing curricula for (and evaluating interventions with) children who are mildly and even moderately handicapped.

Norm-Referenced Assessments

The application of a developmental model using norm-referenced assessments to the severely and profoundly handicapped leads to several problems. The first is that the severely handicapped are not represented in the standardization samples used to construct the developmental schedules. Bayley, for example, in the *Manual for the Bayley Scales of Infant Development* (1969) specifically issues a caveat against use of the instrument for predictive purposes with children who have sensory or

physical handicaps or children who have developmental quotients below 50.

A second problem is the violation of the assumptions underlying the construction of developmental and intelligence tests. Basic among these is the assumption that individuals taking the assessment will have had equal opportunities to learn the information it demands. One has only to meet several severely and multiply handicapped children who at age two or older are not sitting, much less ambulatory, to recognize the frequency with which that assumption is violated when one attempts to utilize such assessments with the severely handicapped.

Several observations offered by Haskett and Bell (1978) in their study of the utility of the *Bayley Scales* with individuals who are profoundly retarded speak to this point. Haskett and Bell suggest that the developmental deficits associated with profound retardation can be described in three ways:

> First, there is an apparent failure of the profoundly retarded to enter into an evolving system of feedback with various sources of stimulation in their environment . . . Second, there is an apparent failure to respond positively—in an emotional or motivational sense—to variations in perceptual-environmental events, and thus a failure to create such variations (which failure might explain the first). Third, there is an apparent failure to develop a symbolic representational system which presumably requires the mediation of other people and which marks the transition from perceptual to conceptual organization of the environment (pages 346-347).

The objects used in scales, such as the *Bayley*, were selected on the basis of their rather universal utility with and interest to infants and toddlers. In general, I have found them quite acceptable reinforcers for the majority of children. With respect to the severely and profoundly retarded, however, I do not feel the materials are quite so appropriate. Their handicaps limit opportunities for interactions with the environment and result in different histories of reinforcement.

An instrument, such as the *Bayley*, can be a rich source of descriptive information about the person who is assessed, for the experienced clinician. Too often however, the only use made of the information is the preparation of a summary score for predictive and classification purposes. Again Haskett and Bell (1978) suggest the limitations of this practice with the profoundly retarded. They point out that the greatest problem they had in assessing their sample was "reliability with respect to behavioral occurrences" (page 347). And they conclude: "On any single administration of the scale, then, the examiner may not be able to secure a representative sample of an individual's behavior, and will, if anything, probably underestimate the repertoire" (page 347).

With respect to the use of the scale as a basis for monitoring the effect of intervention, Haskett and Bell's data suggest "that it may be unwise to assume that any single administration will provide sufficient baseline data against which to evaluate the effects of experimental procedures, at least if the purpose of the analysis is the identification of specific behaviors or specific clusters of behavior that are later modified by intervention" (page 348).

A Dilemma

Conscience and accepted practice in the educational community require that the educator evaluate the effectiveness of interventions with children who have handicaps against appropriate and, preferably, scientifically acceptable developmental criteria. However, those criteria are by their very nature unsuitable for the severely and profoundly retarded. Given that there are no appropriate norm-referenced assessments for this population, I suggest that the essential criterion to be met in selecting a basis for assessment must be the educationally prescriptive utility of the instruments or procedures.

To evaluate the prescriptive utility of an assessment procedure, several concerns must be addressed. These concerns include: What reinforces the child? Are the response criteria appropriate for a given child? Are the instructional objectives to be derived from the assessment information appropriate for a given child? Are the procedures to evaluate learning appropriate for a given child? (See Robinson and Robinson, 1978, for a more extensive discussion of these criteria than is possible in this paper.)

One approach to assessment that we have found useful with severely handicapped infants has been that represented in the *Uzgiris-Hunt Ordinal Scales of Sensorimotor Development.* There are several characteristics of these scales and the approach that was taken in developing them that may account for their utility with children who are severely handicapped. These characteristics are defined by the following questions and answers.

Has the Condition of a Reinforcing Situation Been Met? The scale in the several areas of sensorimotor development (object permanence, means-end and causality, spatial relationships, imitation and schemes in relation to objects) is comprised of "eliciting situations" designed to assess whether a child has use of the sensorimotor schemes as part of his daily interactions with his environment. The use of "eliciting situations," such as putting the child in a situation where he must get an object, out of his reach, with *any available tool* (string, cloth, pillow, box top, etc.), rather than tasks constrained by a specific tool (such as a plastic red ring to be secured with the use of a piece of white plastic string, as is the case in the *Bayley Scales*) is one of the useful characteristics of these scales. This flexibility in materials somewhat solves the problem of ensuring that assessment materials are reinforcing to the

child. The "eliciting situation" approach allows one to try different materials until an object or activity is found which meets the operational definition of a reinforcer.

Are the Response Criteria Appropriate? The "eliciting situations" are designed to help determine whether a child shows evidence of a particular sensorimotor scheme. What is considered an acceptable demonstration of that ability (the response criterion) can be adapted to the response repertoire of an individual child. Thus, the child who cannot physically manipulate a toy in a way to give evidence that he understands causality may be able to convey the same understanding through a gesture. (See Robinson, 1976, Robinson and Robinson, 1978, for more extensive discussions of this point.)

Does the Assessment Aid in Selecting Objectives and Methods? In answer to this question—the approach represented in scales such as the *Uzgiris-Hunt* serves us well. The focus of these assessment strategies is explicitly upon documenting those conditions under which an individual shows evidence of a particular concept or skill. This is in contrast to standardized assessments that typically focus on whether an individual can or cannot perform an activity in a specific manner. An assessment which makes possible the documentation of the conditions, tasks, and response adaptations that elicit a particular form of behavior will be useful in establishing instructional objectives and methods. This approach supports the concepts of the developmental model by providing a description of the individual's current level of functioning in terms of behaviors he can already perform. Thus, it provides the basis for planning the next concepts or skills to be learned.

Emphasis upon Concepts and Strategies

While assessment instruments such as the *Uzgiris-Hunt* certainly illustrate use of the developmental model, what are their advantages, for example, over the *Bayley Scales of Infant Development* in assessment and prescription for the severely and profoundly retarded? A primary advantage, from my point of view, is the direct organization of the *Uzgiris-Hunt Scales* around aspects of sensorimotor development as prerequisites for later cognitive development. This emphasis upon the specific strategies and concepts that constitute, for Piaget, sensorimotor development allows observation of the individual's behavior in various situations, including those that are more typical of his daily environment, as the basis for determining whether in functional situations the sensorimotor schemes or abilities are available to him. The fact that the *Uzgiris-Hunt Scales* are not burdened with the restrictions imposed by a standard set of materials, conditions for administration, and responses obviates the qualifications that the experienced clinician must iterate any time he or she tries to apply a standardized assessment to a situation where it really does not fit, such as with the severely and profoundly retarded.

An illustration of how information derived from the sensorimotor assessment can be utilized in the formulation of individualized instructional objectives is provided at the end of this chapter in the form of an educational evaluation report written by Nancy Fieber. More extensive descriptions of the use of the sensorimotor scales as the basis for designing curriculum for the severely handicapped can be found in Robinson and Robinson (1978).

Evaluation of Effectiveness of Intervention

Elsewhere in this publication writers emphasize the importance of asking the right questions in performing evaluations of program effectiveness. When considering individuals who are severely handicapped, whether with respect to individual assessment or overall program evaluation, the educator must keep in mind the point of asking the "right" questions. The specific short- and long-term goals of the intervention undertaken with the individual who is severely handicapped (and with his family) must be carefully specified. It is unlikely that the severely and multiply-handicapped individual will be able to function in the typical environment without some prosthetics and modifications. To assess such as individual's functioning against criteria established with persons with no significant sensory, physical, or learning problems is not a meaningful activity. And to assess such an individual's family's functioning against the "normal" family's requirements for support and then to make statements about the family's adaptive abilities is also unmeaningful. We have no norms by which to evaluate whether the rate of developmental progress made by one individual who is severely handicapped is consistent with his potential for change. At the present, I feel we can only provide descriptions of changes in level of functioning with respect to the individual's ability to interact with, and operate in, his environment; to do things for himself; or to request aid in those situations where he cannot act for himself. While the form of that action or request cannot be rigidly prescribed, we can set a goal that it approximate as closely as possible the typical form. However, decisions on the criteria for individual persons will have to be clinical for the present time. The clinician will, of course, have the responsibility of providing specific descriptions of the behaviors being accepted as alternatives for the typical form of actions and requests. He will also be responsible for describing the specific behaviors that are the basis of his decisions to select alternative responses.

In summary, the basic tenents for both assessment of individuals and evaluation of the effects of intervention apply to the severely handicapped. The appropriateness of the instrument, in terms of the assumptions made in its development and task and response requirements, must be considered. The criterion against which the assessments should be judged is the usefulness of the information gathered for designing a program of instructional objectives and methods. In selecting ways of evaluating the effectiveness of intervention procedures with individuals who are severely or profoundly retarded, the criterion to be met is: Are the questions being asked in the evaluation applicable to the specific characteristics of the individuals receiving the intervention?

UNIVERSITY OF NEBRASKA MEDICAL CENTER

Reg. No. 27-45-17

MEYER CHILDREN'S REHABILITATION INSTITUTE
444 SOUTH 44TH STREET
OMAHA, NEBRASKA 68131

Name

B'date 12-30-77

EVALUATION X PROGRESS NOTE

Date 9/20/78

History: K has been referred for functional vision and educational assessment. She has been followed in physical therapy and is currently on a home program. Our evaluation was begun on 6/22 but not completed until 9/20/78, because of illness. In addition to a history of seizure disorder and motor delay, she has suspected deficits in both vision and hearing. On 4/3/78, the Pediatric Neurologist noted that she did not fix on faces or shiny objects, and hearing problems were also suspected as she did not startle to sounds. The opthalmologist on 5/8/78 noted that she did not fix or follow well and that an optokinetic nystagmus was not elicited. The impression was reduced vision of central etiology. Reevaluation was recommended in six months. On 5/9/78 a visually evoked EEG was done and showed that light impulses were transmitted to the occipital cortex. Audiologist, 5/10/78, noted some startle response to 4000 Hz at 40 dB sound field. The impression was possible normal peripheral hearing in at least one ear but loss was not ruled out. Impedance testing was not done. Reevaluation in six months was recommended. When seen on 6/5/78 by the Pediatrician, it was noted that she did not fixate objects and hearing was questionable. When seen by the neurologist on 7/26/78, he noted that she seemed to only see things in the far field.

Procedures Used: Piagetian Sensorimotor Cognitive Assessment; functional vision and hearing observations.

1. **General Observations**: Spontaneous behaviors were observed in several positions. In supine, she turned her head to side or midline and engaged in midline hand play. Hands were often at mouth and she vocalized with hands at mouth. She also engaged in light gazing. Placed in prone she pivoted and rolled over, though not yet directed toward object or place.

2. **Functional Vision**:

   A. **Localization and oculomotor behaviors**: One got the impression of a slight internal strabismus because of her wide-spaced eyes, but with pupil reflections this was questionable. At times her eyes appeared cloudy. When first seen on 6/22/78, she seemed to "lock" visually on far stimuli. On 9/20, she seemed better able to look away voluntarily but the tendency to "lock" remained, and it was difficult to obtain fixation on a stimuli presented at close range. She did fixate and track certain stimuli 180 degrees horizontally, some vertically, and circularly with eye-hand movements. She turned to peripheral movements stimulus but did not blink to approaching stimulus or shadow. She did not yet shift gaze between two stimuli without movement and sound cues. When turning to lateral sound she fixated sound sources. She showed no vergent adjustments to stimuli at different distances– when a fixtated object such as face was moved closer there was no convergence and she would break the fixation. On one occasion we saw a visually directed swat, suggesting localization ability.

   B. **Discrimination Behaviors**: In addition to light perception, she was aware of movement at both far and near range. The objects she did fixate and track were fairly large and had shiny light catching qualities, or movement– such as a shiny windmill. She oriented to faces with eye-to-eye contact with voice cues, but has not yet shown visual recognition for her mother without other cues. This may imply form-pattern perception at least at a primitive level of dark spots on light oval, if not more. Visual interest in large red crackly paper could have been shadow or movement at least. There was not yet confirming evidence of color, defined form and pattern vision, but this does not mean they are not there. The face interest and tendency to be caught by far stimuli suggest something beyond gross light-shadow perception and distance. Probably because she is not adjusting focus for different distances things could be blurry, like a camera out of focus.

   Implications for the present are to continue to encourage looking behaviors, to use stimuli which are shiny, have movement and sound and to provide contrasting background such as a dark visual surface and screen to shut out extraneous stimuli. Reinforce visual tracking but tie it to meaningful sounds, handling toy, etc. Work on shift gaze might start using of sound as cues. Whenever she is visually interested in something, move it closer to challenge her to try to adjust her focus.

3. **Functional Hearing**: Often she does not respond to sound because she was visually "locked" on something. We did see some attending to sounds such as musical toys, and she would at times localize and turn to sound source presented laterally 6-12 inches from her ear. She was aware of presence of voice and attended to it. Goals should be in area of increased attending with clear evidence of awareness of presence versus no presence of sound, and interest in changing sound for example, varied animal sounds of See'N'Say.

4 Cognition

Means for achieving environmental events, causality and object schemes

She engaged in some hand watching at arm's length and was just beginning to look at midline hand play, important precursors of visually directed reach and grasp. She had a voluntary grasp scheme (deliberately trying to maintain) but had difficulty maintaining grasp more than a few seconds due to motor problems. Grasp occurred on contact and not foresightfully with reach as a means to an end. Coordination of grasp with other schemes was explored. When a bell was tied in her grasp she brought it close to her face and moved it across her eyes. When assisted to bring a grasped object to her mouth, active mouth movements occurred, but there was no effort on her part to repeat bringing the bell to her mouth even in sidelying. She has, however, mouthed her wrist. Although she moved arm with bell in grasp in supine, we were not sure this was with awareness of sound. Repetion of movement producing a sound-visual spectacle was explored in several ways.

She repeated swatting movements with hand open in supine, the right arm and hand available for reaching seemed best. In swatting, the movement of the slinky seemed visually interesting. As she swatted the brass windchimes she looked, and it was not certain whether it was the visual movement, its shinyness or its sound which was interesting . . . Some beginning visual directing of approach was seen emerging when chimes were presented first right, then left, and she attempted to reach with right over to left. Grasping occurred on contact after swatting a few times, as if an after-thought.

Instruction needs to emphasize coordination of grasp with looking, mouthing, noise-making shaking as well as working on simply grasping longer. One recommended strategy is to tie object in grasp and assist her to produce single schemes. Following typing, several opportunities to hold should be tried. Swatting visual-auditory spectacles needs to be generalized across variety of objects with careful consideration of stimulus properties that have been demonstrated to be interesting to her. A next step will be foresightful beginning swat or approach on seeing stimulus when placed near hand but not coming into physical contact.

Vocal Imitation K engaged in vocal play, showing attitudinal differences (sounds heard were ah, yeh, eh, da, ga, g, n). Her mother reported she might repeat sounds when her mother imitated her. This would be good to continue to encourage.

5 **Communication - Social** We were interested in signal behaviors she might show and signal functions, as well as what might elicit recognition or anticipation responses.

K soothed preferentially to her mother's handling, and seemed to know her mother through handling, possibly voice. Some nice signs of recognition and enjoyment of an auditory-motor game were seen in a kissing game her mother demonstrated. Smiles were observed in the personal function--expressing pleasure, but not yet in the interactional function. Eye contact was given by K to her mother as her mother talked to her. K showed a nice anticipatory response when given touch plus verbal cues on being picked up. At this time her mother was not aware of other things or events which might be acting as signifiers to K.

Communication goals should be to look for increased signs of recognition, then anticipation, within the context of structured multisensory auditory-motor baby games. We also want to see more situtations which she anticipates when systematic touch plus verbal cues are given. These should precede all handling. They should be systematically developed as integral parts of the way she is handled. Each person who deals with her should create a "name sign" for themselves--a special verbal-touch greeting. A responsive interactional smile would be an appropriate goal--because of her visual problems it will be important to use *recognizable* touch input as one talks to her. A game structure is a good way to do this. When a game becomes familiar and she seems to keep on thinking about it a few seconds when paused, begin trying to elicit a signal for "more whatever."

Recommendations

1. Participation in the Infant Development Program with monitoring and reevaluation of possible multisensory impairments. Methods appropriate for children with combined sensory impairments are recommended.

2. Reevaluation in ophthalmology and audiology in November as previously recommended.

Nancy Fieber, M.A.
Infant Development Program

Cordelia C. Robinson, Ph.D.
Director, Infant Development Program

References

Bayley, N. *Manual for the Bayley Scales of Infant Development.* New York: Psychological Corporation, 1969.

Haskett, J., and Bell, J. "Profound Developmental Retardation: Descriptive and Theoretical Utility of the Bayley Mental Scale." In C. E. Meyer (Ed.), *Quality of Life in Severely and Profoundly Mentally Retarded People: Research Foundation for Improvement.* Washington, D.C.: American Association on Mental Deficiency, 1978.

Friedlander, B. Z., Sterritt, M., and Kirk, G. E. (Eds.), *Exceptional Infant.* (Volume 3) *Assessment and Intervention.* New York: Brunner/Mazel, 1975.

Robinson, C. C. "Application of Piagetian Sensorimotor Concepts to Assessment and Curriculum for Severely Handicapped Children." *American Association for Education of Severely/Profoundly Handicapped Review* 1 (1976): 5-10.

Robinson, C. C., and Robinson, J. H. "Sensorimotor Functions and Cognitive Development." In M. Snell (Ed.), *Systematic Instruction of the Severely and Profoundly Handicapped.* Columbus: Charles E. Merrill, 1978.

Tjossem, T. D. (Ed.) *Intervention Strategies for High Risk Infants and Young Children.* Baltimore: University Park Press, 1976.

Uzgiris, I., and Hunt, J. McV. *Assessment in Infancy: Ordinal Scales of Psychological Development.* Urbana, Illinois: University of Illinois Press, 1975.

CHAPTER 8

Severely and Profoundly Handicapped: How do you measure communication skills?

by Rebecca Fewell DuBose

The ability to communicate to another person is a critical life-skill for all human beings. We begin sending messages with our first cries, and we gradually develop and refine our communications during our later years. Severely handicapped children may never communicate using the symbol system of those around them, but they may develop a signal system that permits communicative exchanges that are recognized and responded to by those who care for them.

In the assessment of, and programming for communication in severely handicapped children, it is helpful to understand the developmental sequence of both receptive and expressive communication and the interactions that signal communicative intent.

Signal and Symbolic Communication

Two levels of expressive language are identifiable in early child language: signal and symbolic level language. Signal behavior is not uniquely human, rather it is shared by many members of the animate world. It occurs when a message is sent by one source and is interpreted by another. The signal represents a physical response to an immediate perceptual event. Lights, buzzers, horns, babies, animals and insects are all signal sources that frequently emit readily identifiable intent.

Symbolic language is characterized by one's use of words or signs that signify certain thoughts, ideas, reactions, experiences, events and places that are not in the immediate perceptual experience of the child. Symbols do not lead to the satisfaction of immediate physical pleasure, but are things to be contemplated. A word becomes a symbol for a child when he uses it in the same way that others in his world use it, and he is obviously letting the word represent something that is not immediately tangible (DuBose, 1978).

Cognitive and Social Communication. Communication derives from experiences that signify our mental awareness of our surroundings and from our desires to interact with persons within our environment. Researchers have approached early communication by examining sensorimotor development and early patterns of human interaction. Both approaches are valuable in identifying early signal and presymbolic indicators of communicative behavior. It is within these exchanges that one finds many of the forms of communication that are used by severely and profoundly impaired children.

Communication in the Sensorimotor Period. Dunst (1978) has compared Piaget's stages of sensorimotor development to Bates' proposed sequence of communication. Additionally, Dunst provides examples of typical behavior during each stage, suggests the child's communicative intent, and the adult's interpretation of the behavior. Table 1 presents the Dunst analysis.

TABLE 1

Response Model Sequenced by Level of Difficulty

1. The testee *affirms* or *negates* the selection of the response by the clinician.
2. The testee *sorts* (pictures, objects, forms, sounds) according to categories which correspond to *models* presented by the clinician. (Categories may be size, shape, color, pitch, texture, or function.)
3. The testee *compares* and *selects* or *points* to a picture, object, form, or letter to correspond to a *model* presented by the clinician.
4. The testee sorts as in (2) but without a model.
5. The testee *matches* pictures, objects, forms, or sounds which are alike.
6. The testee *selects* one of a series which is *different* from the rest in the series.
7. The testee *arranges* in *series* according to gradation in size, color, or some other standard.
8. The testee *verbalizes* responses to any of the above.

Carrow, E. "Assessment of Speech and Language in Children." In *Language Intervention With the Retarded: Developmental Strategies*, edited by J. E. McLean, D. E. Yoder, and R. L. Schiefelbusch. Baltimore: University Park Press, 1972.

Table 2

INNER LANGUAGE SCALE

	Object Concept	Object Permanence	Causality	Imitation
1-4 Months	Child follows an object or light with eyes through 180° arc. Eyes linger at spot object or light is last seen.	Child follows an object or light with eyes through 180° arc. Eyes linger at spot object or light is last seen.	Child engages in hand-watching behavior.	If adult mimics child's arm or hand movements, child imitates adult in turn.
4-8 Months	Child engages in simple motor acts with an object, pauses, then resumes same actions.	When an object is removed, child searches briefly for it. Search is confined to one modality.	When an object is in the visual presence of the child, the child engages in simple motor acts previously experienced in continuity with the object in an apparent effort to maintain or reproduce the activity of the object.	Using actions in his repertoire, child imitates adult's hand or arm movements. (Note: Here the adult initiates the action whereas in No. 4, the child initiated the action.)
8-12 Months	Child intentionally manipulates objects in an exploratory manner, turning them, feeling surfaces, etc. (Note: Here the child is interested in the object itself; in No. 5 child's interest is mainly in the motor acts he performs with the object.)	When an object is taken from a child, he searches persistently at the place from which the object was removed, even though he saw an adult place it elsewhere. (Note: Search will be multisensory.)	Child waits for adult to do things for him; attacks a barrier to a toy as if it caused him frustration. (Note: Here the child is able to preceive others as causes of of activity; in No. 7 he was aware mainly of himself in connection with activity.)	Child imitates adult's actions using actions in his repertoire. He can now imitate actions that he cannot see his body perform.
12-18 Months	Child drops an object repeatedly and follows the fall of the object through space; this is performed with several different objects.	When an object is hidden, child will search for it in the last place he observes it being hidden. (Note: In No. 10 the child cannot follow the object to a new locale; in No. 14 he can follow the object to a new locale provided he observes it being hidden.)	Child signifies desire for adult to repeat an action by returning object to adult and waiting.	Child can imitate actions new to his repertoire through gradual approximations; child imitates action in a more precise manner.
18-24 Months	Child's play with toy facsimiles of objects (comb, cup, spoon, etc.) demonstrates awareness of adult use of the object.	Child searches for lost or hidden objects he may not have seen for several days. (Child does not have to observe the object being hidden.)	Child attempts to re-activate an object he observed in action, or discover the cause of the action.	Child correctly imitates new actions even after a time lapse.

TABLE 3

Name of test: Sequenced Inventory of Communication Development (SICD)
Author: Dona L. Hedrick, Elizabeth M. Prather, Annette R. Tobin
Where test can be purchased: University of Washington Press, University of Washington
Seattle, Washington
Cost: $12.00
Developmental areas/skills covered: Receptive and Expressive Language
Age range/handicapping conditions: 4-48 months mentally retarded, developmentally delayed, learning disabled.
Standardization: 252 children, 21 at each of 12 age levels.
Reliability and validity data: Interexaminer reliability reported on 16 subjects tested by two independent examiners. Mean agreement was .96. Test - retest reliability on 10 subjects tested one week later with 92.8 percent agreement. Correlation of Receptive Communication age to Peabody Picture Vocabulary Test was found to be .8097. Expressive Communication Age and Mean Length Response correlation was .76 and between Expressive Communication Age and Structural Complexity Score was .74.
Who can administer: Clinicians
Training needed: Clinical language training
Comments: (short discussion on philosophical/theoretical approach behind instrument and/or strengths and weakness)

The SICD was developed for clinicians to assist them in remedial programming for the deviant child. It provides an estimated level of development in the areas of receptive and expressive communication. The data provides information concerning what a child responds to or fails to respond to. From this diagnostic data a teaching program can be designed.

The normative sample is small for the age span; more research is needed.

Name of test: Environmental Language Inventory (ELI)
Author: James D. MacDonald, Marjorie Nickols
Where test can be purchased: Charles E. Merrill
1300 Alum Creek Drive, Box 508
Columbus, Ohio 43216
Cost: $6.75
Developmental areas/skills covered: Expressive language
Age range/handicapping conditions: 24-54 months, mentally retarded; developmentally delayed
Standardization: 25 normal preschool children, 5 subjects at each of 5 ages levels.
Reliability and validity data: ELI responses and free speech samples were scored in terms of semantic-grammatical rules and mean length utterance. Interjudge reliability was 97 and 98 percent agreement. Data presented in the manual supports finds of earlier researchers such as Brown, Bloom and Schelsinger, as to the age at which semantic rules appear.
Who can administer: Clinicians with a training in speech and language development.
Training needed: Clinical language training
Comments: (short discussion on philosophical/theoretical approach behind instrument and/or strengths and weakness)

The ELI is a semantic-based measure for the assessment of early expressive language. It represents an approach to diagnosis and training of language through analysis of what a child spontaneously says or is prompted to say through conversation, or imitation. The analysis of utterances permits one to program for grammatical rules that are not currently used by the child.

The ELI's standardization sample is extremely small. More information on reliability and validity is needed.

TABLE 4
Curricula Sources for Communication Training

Source	Format	Comment
"Training Prerequisites to Verbal Behavior," D. Bricker and L. Dennison in M. E. Snell (Ed.), *Systematic Instruction of the Moderately and Severely Handicapped*, Columbus: Charles E. Merrill, 1978.	Book Chapter	Emphasis is placed on a task analysis of the prerequisites to verbal language. Four training areas are presented: 1) on-task behavior; 2) limitation; 3) discriminative use of objects; and 4) word recognition.
Language Acquisition Program for the Severely Retarded. L. Kent. Champaign, IL: Research Press, 1974.	Book	A behavioral program that contains a highly systematic receptive phase. Includes many language games in the teaching strategies.
Ready Set Go - Talk to Me. D. Horstmeier, J. D. MacDonald and Y. Gillette (Eds.), Columbus: Charles E. Merrill, 1978.		Instructional program for early language acquisition.
"A Syntax Teaching Program," J. F. Miller and D. E. Yoder in J. McLean, D. Yoder and R. Schiefelbusch (Eds.), *Langauge Intervention with the Retarded*. Baltimore: University Park Press, 1972.	Book Chapter	The strategies of Miller and Yoder stress a semantic-based intervention program. The strategies can be implemented once the child is using a few simple words.
and "An Ontogenetic Language Teaching Strategy for Retarded Children," J. F. Miller and D. E. Yoder, In R. L. Schiefelbusch and L. L. Lloyd (Eds.), *Language Perspectives: Acquisition, Retardation and Intervention*. Baltimore: University Park Press, 1974.	Book Chapter	A sound basis for language training is provided noting the relationship between form, function, and experience.
"A Behavioral-Psycholinguistic Approach to Language Training," in L. McReynolds (Ed.), *Developing Systematic Procedures for Training Children's Language*. American Speech and Hearing Association Monograph 18 (1974): 96-124.	Book Chapter	Program is based on a syntax approach to skill acquisition.
BKR Educational Projects, Inc., L. M. Bradtke, W. J. Kirkpatrick, Jr., K. P. Rosenblatt. BKR Educational Projects, Inc., 1970 S.W. 43rd Way, Fort Lauderdale, FL 33317.	Kit	Fifteen basic language lessons included in a program that contains sections on other behaviors. Language emphasis on receptive, expressive, nonverbal communication by sign. Activity card format.
Karnes Early Language Activities, Author: Merle Karnes, GEM P. O. Box 2339, Station A, Champaign, IL 61820.	Kit	This program is a downward extension of the GOAL Program, but is not sequential. It trains receptive, expressive, and motor imitation. There are 200 model lessons and 1,000 activities on card.
Merkley Developmental Approach to Speech and Language. Author: Frances A. Merkley, Project A.D.A.P.T., Tri-Counties Regional Center, 22 West Michetorena, Santa Barbara, CA 93101.	Kit	Targeted for all developmentally disabled persons, including the deaf, ages 2-8 years. This kit contains programming for receptive, expressive, motor imitation, and nonverbal communication.
Monterey Language Program (Programmed Conditioning for Language). Authors: Burl R. Gray, Bruce P. Ryan, Monterey Learning Systems, 900 Welch Road, Palo Alto, CA 94304.	Program	The 40 lessons contain twenty-five steps each with screening test, pretest and posttest. Users must be trained by a certified trainer. The program is highly structured but offers maximal opportunity for prescriptive programming. Receptive and expressive language are programmed.
Non-Speech Language Initiation Program, Joseph Carrier, Tim Peak, H & H Enterprises, Inc., Box 3342, Lawrence, KS 66044.	Kit	This program is designed to teach the non speaking retarded person to generate seven-word spoken sentences. The kit contains 188 structured lessons with tests. Expressive, receptive, motor imitation, and nonverbal communication by plastic words (abstract symbols) are taught.

Perlocutions are signals that are unintentional, but do have the effect of making the onlooker feel as though the child was communicating. Neither the sender or receiver recognizes the intent as a conventional form of communication. This stage occurs from birth to about ten months of age. *Illocutions* are conventionally recognized, nonverbal signals intentionally used to gain the attention of another person, make a request, etc. This stage ranges from ten to between twelve and fifteen months. *Locutions* represent combinations of conventional verbal utterances (words) in meaningful ways to serve the same function as illocutionary acts. The locutionary stage begins around twelve to fifteen months and goes to twenty-four months (Bates, 1976).

Using a similar paradigm, but one viewing the cognitive aspects of communication, Branston and DuBose (1974) developed the "Inner Language Scale" which permits the examination of interactions between the child, objects, and people, using Piagetian schema. This model has proven helpful in examining the communication of this population.

Assessment of Verbal Communication. Severely handicapped children respond best in environments that are familiar to them and when they are exposed to objects and events that are likewise a part of their surroundings. For this reason, assessment measures that permit the child to engage in very common tasks and do not require responses to a standard set of stimuli are needed. The two tests presented in Table 3 have been found to be particularly amenable to the assessment of these children.

Results from these tests (Table 3) are readily translated into educational programs. Both the cognitive and social aspects of language can be fostered through content programmed from these results.

Curricula for Communication Development. Several curricula and other educational materials are available and are helpful in planning for communication development. Those noted in Table 4 address at least a portion of their content to some aspect of communication that is appropriate in severely impaired children.

Communication in severely handicapped children can best be examined in spontaneous environments or in those environments where the child's expressive acts are most likely to be emitted. As should naturally follow, it is equally as important to train a child in language in natural environments. While doing so, it is critical that the teacher or parent systematically instruct and monitor the child's progress. Once responses have become a part of the child's repertoire, the need to generalize the responses becomes a major goal. Brown, Nietupski, and Hamre-Nietupski (1976) have carefully stated the extent to which this should occur: "Each time a student is taught a functional skill or game, it might be required that the skill be performed: (a) in reaction to, or in the presence of at least three different persons; (b) in at least three different natural settings; (c) in response to at least three different sets of instructional materials; and (d) to at least three different appropriate language cues" (p. 10).

The procedures for assessment and training presented are based on the philosophy that communication occurs when the child has a desire to express something, and that persons in his world are receptive to his expressions and will respond to him, thus reinforcing his communicative efforts and shaping them into more acceptable responsive forms. The procedures represent an attempt to help the child initiate his own efforts, rather than training him to respond in ways that others have dictated for him.

References

Bates, E. "Pragmatics and Sociolinguistics in Child Language." In *Normal and Deficient Child Language*, edited by D. Morehead and A. Morehead. Baltimore: University Park Press, 1976.

Branston, M. B., and DuBose, R. F. *Inner Language Scale*. Duplicated paper, George Peabody College for Teachers, Nashville, Tennessee, 1974.

Brown, L., Nietupski, J., and Hamre-Nietupski, S. "The Criterion of Ultimate Functioning and Public School Services for Severely Handicapped Students." In *Hey Don't Forget About Me: Education's Investment in the Severely, Profoundly and Multiply Handicapped*, edited by M. A. Thomas. Reston, VA: Council for Exceptional Children, 1976.

DuBose, R. F. "Development of Communication in Nonverbal Children." *Education and Training of the Mentally Retarded* 13 (1978): 37-41.

Dunst, C. "A Cognitive Social Approach for Assessment of Early Nonverbal Communicative Behavior." *Journal of Childhood Communication Disorders* 2 (1978): 110-123.

SECTION 3: OTHER ISSUES

Editor's Notes

Three of the presentations at the conference dealt with issues of a somewhat different nature. Rather than focusing on specific instruments for early childhood programs, these authors discussed more global issues which must be considered when selecting and using instruments.

The issues in chapter nine relate to the implications of PL 94-142 on testing, specifically that portion of the law dealing with nondiscriminatory testing. Harbin presents each provision in the *Protection of Evaluation Procedures* section of the Law along with its particular implications for testing. Several broader implications of nondiscriminatory testing are also discussed including: redefining roles, future training, developing new instruments, and the acceptance of the idea of cultural pluralism. Finally, Harbin outlines a model for demonstrating compliance with the mandates of the law. It includes checklists that can help identify compliance questions and how to document responses.

The reader should be aware that the term "evaluation" in PL 94-142 refers to any type of examination of a child, e.g., screening, diagnosis, and assessment, rather than to the examination of the total program. Also, the reader should note that this chapter deals only with the implications of the *Protection in Evaluation Procedures* portion of the Law. Other parts of PL 94-142 may also have implications for testing, but are not addressed in this discussion.

Chapter ten addresses the need for and the problems in adapting tests and testing procedures to accommodate the uniquenesses of the young handicapped child. DuBose acknowledges a "Catch 22" dilemma: if you administer a standardized test maintaining strict adherence to its procedures, you know the handicapped child will not be able to perform the items; on the other hand, if you vary the procedures so that the child can perform the items, you know the test is no longer standardized. She provides a strong rationale for the necessity of adapting tests and testing situtations if they are to provide meaningful information, and she offers some suggestions on how to deal with the resulting problems.

Chapter eleven focuses on one particular aspect of program evaluation, the task of choosing data collection designs and instruments that are compatable with each other. First, Sheehan defines *instrument* to include more than just commercial standardized tests. He then presents five types of data collection designs. Each is basically quantitative, i.e., it seeks to provide evidence of child progress using scores on instruments. In order to argue that the progress is the result of intervention, the outcomes are compared to results achieved without intervention. Finally, he suggests a decision-making model for selecting a design and for selecting instruments that will fit the chosen design. He cautions that each of the designs presented have problems when applied to young handicapped children.

There are certainly more measurement issues for the field than can be discussed in this publication, such as establishing test norms for handicapped populations or the use of case studies as a qualitative data collection design to name two. It is hoped that the issues presented here will stimulate further thinking and discussion among HCEEP staffs and thereby improve measurement practices for the benefit of young handicapped children.

CHAPTER 9
Public Law 94-142: What are its implications for measurement?
by Gloria Harbin

Juan is a four-year-old Mexican-American child. His parents do not agree with the decision made by his school to place him into a program for developmentally delayed children. They are contesting the decision. The school psychologist gave Juan the WIPPSI, and states that Juan received a full scale score of 70. The school also says that Juan did not get along with the other children his age, did not follow directions, and was unable to participate in most of the activity lessons. They feel that in comparison to children his own age, Juan is delayed and needs a special program. Juan's parents feel that he is not retarded and, if he received some help in the regular class with English, he would progress faster. Since Juan is bilingual, it is often difficult for him to understand just what is happening in the classroom.

This is just one example of a number of hearings being held around the country. They occur because people who test handicapped children do not comply with the nondiscriminatory testing mandates of PL 94-142. They use tests and/or procedures which discriminate on the basis of race, culture, or handicap: the result being that inaccurate or distorted information is used to make placement decisions.

The Protection in Evaluation section of PL 94-142 was written in order to stop inappropriate labeling and placement of minority and handicapped children. In the late 1960's and early 1970's, several court cases demonstrated injustices to children as a result of poor testing procedures (Aureola v. Santa Ana Board of Education, 1968; Spangler v. Pasadena Board of Education, 1970; Diana v. California State Board of Education, 1970; Larry P. v. Riles, 1971). These cases brought out major criticisms of testing practices, for example: (1) many tests placed a heavy emphasis on verbal skills; (2) some test questions were culturally biased; (3) testing procedures ignored the child's ability to learn in Spanish or any other language; (4) most tests had been standardized on white, middle-class children; (5) parents were not adequately informed and involved in the placement process; (6) intelligence tests do not adequately measure the learning abilities of minority group children. PL 94-142 deals with the criticisms by setting down legal standards for the selection and use of testing materials and procedures. The intention is to eliminate the biases against children with handicaps and children from different racial and cultural groups.

In this chapter, each of the law's mandates concerning the testing of young handicapped children will be examined–individually. Then we will review the major implications of the mandates for the educational community, and conclude with a model that child workers can use in attempting to comply with the requirements. The purpose of the model is to assist professionals in systematically identifying and minimizing biases in their testing procedures.

The Mandates
(Final Regulations for PL 94-142: Protection in Evaluation Procedures)

†121a.530 General.

(a) Each State educational agency shall insure that each public agency establishes and implements procedures which meet the requirements of ††121a.530-121a.534.

(b) Testing and evaluation materials and procedures used for the purposes of evaluation and placement of handicapped children must be selected and administered so as not to be racially or culturally discriminatory.

Regardless of their agency (education, mental health, social services, etc), all professionals who provide information to be used in a placement decision are responsible for following the requirements of this law. Furthermore, in selecting a test for a particular child, the composition of the group of children on which the test was normed must be examined to determine if children similar to your child were included in the sample. For example, it is racially and culturally discriminatory to administer a test normed on white middle-class children to an Afro- or Mexican-American child and assume that his test performance is an accurate reflection of his standing in the total population. It is not. It is a comparison of his performance with the performance of white middle-class children only.

†121a.531 Preplacement evaluation.

Before any action is taken with respect to the initial placement of a handicapped child in a special education program, a full and individual evaluation of the child's educational needs must be conducted in accordance with the requirements of †121a.532.

There have been numerous instances where a regular classroom teacher decides that a child in her class needs help and recommends placement in a special education class. The principal tells the special education teacher that the child is failing in the regular class and his only hope for success is in her class. The child is then placed in the special education class *without* the appropriate testing or permission. Only those children who have full educational assessments, meet eligibility requirements, and have parental permission can receive special services.

†121a.532 Evaluation procedures.

State and local education agencies shall insure at a minimum, that:

(a) Test and other evaluation materials:

(1) are provided and administered in the child's native language or other mode of communication, unless it is clearly not feasible to do so;

Children must be assessed in their dominant language. While this is usually standard English, it may also be Spanish, Vietnamese, etc. Children who use a mode of communication other than oral (e.g., manual communication or total communication), must be tested in their primary mode of communication. Compliance with this rule will be difficult at times. If there is no one on your staff who can administer the test in this language or mode of communication, you may use an interpreter.

(2) Have been validated for the specific purpose for which they are used; and

The term "appropriately validated" has been used by many professionals to explain this particular mandate (Brantley, 1978). There are basically three purposes for tests: (1) screening; (2) placement, classification, or diagnosis; and (3) educational programming. As a result of this mandate, tests must be used only for the purpose for which they were validated. For example, the Denver Developmental Screening Test can be used only for screening and not for classification. The term "appropriately validated" also encourages users to examine the quality of the validity, reliability, and the normative population for each test under consideration. If the validity and reliability of the test are low, then the assumptions which can be made based upon the child's performance are limited.

(3) Are administered by trained personnel in conformance with the instructions provided by their producer;

The administration process cannot be changed on the basis of professional judgment. For example, the person giving the test cannot change wording, time limits, or order. Some professionals feel that rigid adherence to this mandate may actually discriminate against some minority children and certain handicapped children, in addition to contradicting †121a532(c). In such cases it is recommended by these professionals that the test be administered first according to the directions and then, if necessary, modified. Record both scores and be sure to indicate in the report the modifications that were used as well as the score from the modified administration. The alterations in the test must be justifiable based upon the characteristics of the child.

In some school systems with few school psychologists, speech clinicians, etc., teachers used to be given the responsibility of administering tests without being adequately trained to do so. The scores from these tests were then used as valid scores. This practice is no longer permitted.

(b) Tests and other evaluation materials include those tailored to assess specific areas of educational need and not merely those which are designed to provide a single general intelligence quotient.

Placement decisions based primarily upon information obtained from intelligence tests are no longer permitted. This will drastically change the assessment process in many school systems and programs. Tests which are relevant to the child's educational program must also be given to pinpoint the child's strengths and weaknesses more precisely. This mandate is likely to change the roles and skills of educational professionals, such as the school psychologist and the classroom teacher. One of its aims is to provide relevant information to those who are developing the child's Individualized Education Program (IEP). For years, teachers have complained that information from intelligence tests is of little help in planning and selecting educational experiences for children. This mandate deals with that criticism by requiring the use of educationally relevant tests and materials.

(c) Tests are selected and administered so as best to ensure that when a test is administered to a child with impaired sensory, manual, or speaking skills the test results accurately reflect the child's aptitude or

achievement level or whatever other factors the test purports to measure, rather than reflecting the child's impaired sensory, manual, or speaking skills (except where those skills are the factors which the test purports to measure).

The implications of the mandate above are staggering. Assessment devices must match the characteristics of the child, especially in the case of hearing impaired, visually impaired, physically impaired, and linguistically impaired children. For example, a test requiring motor responses to items which measure cognitive ability cannot be used with physically handicapped children. Strict compliance with this section will be difficult due to the characteristics of instruments that are currently available.

(d) No single procedure is used as the sole criterion for determining an appropriate educational program for a child; and

(e) The evaluation is made by a multidisciplinary team or group of persons, including at least one teacher or other specialist with knowledge in the area of suspected disability.

The practice of one professional administering one device and making the placement decision is no longer permitted. A group of professionals possessing expertise in different areas must administer a number of educationally relevant devices which can include observation and criterion-referenced measures. The multidisciplinary team should include a specialist in the area of the child's disability and a teacher who knows the child, as well as other needed professionals.

(f) The child is assessed in all areas related to the suspected disability, including, where appropriate, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status, and motor abilities.

Comment. Children who have a speech impairment as their primary handicap may not need a complete battery of assessments (e.g., psychological, physical or adaptive behavior). However, a qualified speech-language pathologist would (1) evaluate each speech impaired child using procedures that are appropriate for the diagnosis and appraisal of speech and language disorders, and (2) where necessary, make referrals for additional assessments needed to make an appropriate placement decision.

All children suspected of needing special education services will now receive more comprehensive assessments which will present a more complete picture of each child. Furthermore those instances in which children were referred, for example, because of suspected language disabilities and were given an intelligence test and, perhaps, achievement tests, but were never tested in the area of suspected disability, will stop.

†121a.533 Placement procedures.

(a) In interpreting evaluation data and in making placement decisions, each public agency shall:

(1) Draw upon information from a variety of sources, including aptitude and achievement tests, teacher recommendations, physical condition, social or cultural background, and adaptive behavior;

The multi-disciplinary team must use a variety of sources for placement besides information from educational tests. The child's physical condition, cultural background, and adaptive behavior must all be considered. Adaptive behavior is defined by the Office of Civil Rights as the degree to which the child is able to function and participate effectively as a responsible member of his family and community. At this time, instruments which measure adaptive behavior and sociocultural levels of preschool children and which also meet the other relevant mandates of PL 94-142 (adequately normed and educationally relevant) are nonexistent. Thus, strict compliance will be difficult. However the intent of the law can be satisfied by asking parents if the child is able to accomplish at home tasks specified in norm-referenced developmental devices (e.g., Developmental Profile by Alpern-Boll).

(2) Insure that information obtained from all of these sources is documented and carefully considered;

The multi-disciplinary team is now accountable for gathering and documenting data used to make placements. When such information is documented, it is more difficult for professionals to make hasty placement decisions (inappropriate labeling) which could have adverse effects on children.

(3) Insure that the placement decision is made by a group of persons, including persons knowledgeable about the child, the meaning of the

evaluation data, and the placement options; and
(4) Insure that the placement decision is made in conformity with the least restrictive environment rules in ††121a.550-121a.554.

Compliance with this mandate requires highly skilled professionals be given adequate time in their schedules for assessment activities and multi-disciplinary team meetings and that such people as parents and teachers also be a part of the process since they usually know the child best. Furthermore, each setting should have a continuum of services which is known to members of the multi-disciplinary team so they can select the least restrictive environment for the child. So much confusion and misinformation surrounds the definition of "least restrictive environment" and what it means for individual children that it is no wonder many professionals are baffled and have a difficult time in making this kind of decision. This confusion is compounded by the fact that many school systems do not have the continuum of services needed by preschool handicapped children.

(b) If a determination is made that a child is handicapped and needs special education and related services, an individualized education program must be developed for the child in accordance with ††121a. 340-121a.349 of Subpart C.

The child's IEP must be based upon comprehensive information about the child's present level of functioning in all areas assessed, the child's strengths and weaknesses, and factors such as sociocultural level, previous experiences, and other factors which influence learning characteristics. Thus, the development of the IEP must be systematic and begin with a multi-faceted, multi-source assessment by a multi-disciplinary team.

As you can see, in Juan's case many of the mandates were not followed: (1) Juan was not tested in both Spanish and English. (2) He was not tested by a multi-disciplinary team. (3) The placement decision was based primarily upon an intelligence test. (4) There was no real use of educationally relevant tests. (5) Juan's assessment was not comprehensive. (6) It did not include information from a variety of sources. (7) There was no measure of adaptive behavior. As in Juan's case, the failure to adhere to one or several of the mandates can result in a distorted or biased picture of the child, with civil rights violated and the probability of misclassification increased.

Educational Implications

The impact that this law will have on the field of special education in the coming years will be multifaceted. In this section, we will examine the scope of some of the major changes it will necessitate in the immediate future. The areas that we are reviewing here are by no means the only areas that will be affected profoundly by PL 94-142.

1. **Legal Standards.** PL 94-142 establishes, for the first time, a set of *legal* standards for testing children. Prior to this, the American Psychological Association developed standards titled *Standards for Educational and Psychological Tests* (1974). Unfortunately, many professionals did not follow these guidelines for several reasons: 1) many did not know the standards existed; 2) many were not properly trained; 3) and others did not understand the rationale behind the standards and thus were indifferent about following them. Consequently, much time was spent in trying to convince others to follow these standards in test development, selection, and administration. In addition to establishing *legal* standards for professionals, PL 94-142 also provides a vehicle for legal recourse and legal remedies when a child's parent or advocate feels that the assessment and placement process was discriminatory, distorted, or biased.

2. **Individualization.** Testing procedures must now be individualized matching the child's characteristics with the characteristics of the test. One test may be right for one child but totally inappropriate for another. Giving a battery of tests to every child regardless of handicapping condition, cultural background, and language ability is now obsolete. Professionals must learn the strengths and limitations of assessment devices and become familiar with a variety of instruments to meet the needs of children with diverse characteristics.

3. **Test Selection and Administration.** The name or popularity of a test can no longer guide selection. Now, the following must be considered: the norm group (its size and composition), the information on test validity and reliability, the language of the test, and the modalities needed by the child to *understand* and *respond* to the items. Test publishers will probably take these factors into account when developing new tests and restandardizing old ones.

4. **New Instrument Development.** Developers must now specifically state the purpose of the test. They can no longer claim that they have a (a) screening device that (b) provides diagnostic information which (c) can be used for daily planning. To make such a claim means that the test must be validated for these three purposes, which is psychometrically almost impossible.

Developers and publishers now must also pay more attention to the size and composition of the norm group. Consequently, new tests will probably include children from minority groups and different regions of the country more frequently in their norm population. This will take more time and money. Also many assessment devices currently in use are unacceptable because of the new standards, thus some educational areas now have no appropriate devices. New instruments desperately need to be developed, especially for children with sensory and physical handicaps.

5. **Redefinition of Roles**. Many professionals working with young handicapped children must now play their professional roles differently. In the past, a psychologist usually tested the child and decided on placement. Now that others, including parents, must be included in the process, roles will have to be redefined and new skills learned. All must learn to trust and value the information gathered and presented by others. The skills used by professionals who work as individuals are very different from those needed to work as a team. Programs for the preschool handicapped child fortunately are much farther along than programs for school-age children in their acceptance and use of the multi-disciplinary team approach. Considerable room for improvement exists across the country, however, and we must continue to provide leadership in helping others function effectively as members of the teams.

6. **Future Pre- and Inservice Training**. Pre- and inservice training programs must now focus on helping professionals develop skills in their own areas of expertise while also helping them develop the skills and attitudes necessary for working as team members. Inservice training is needed to help those already in the field perform tasks for which they have had little or no training. As a result of mainstreaming, many regular classroom teachers have been asked to work with young handicapped children, which requires knowledge of individualization, special techniques, and educational materials for different types and levels of children. Psychologists, who are used to testing older children, need to become more familiar with tests for preschool children as well as the techniques needed in testing them. Preservice training, such as in university programs, is going to have to change focus and emphasis to prepare professionals for the "real world." For example, if people are to work as members of multi-disciplinary teams they must not be trained in isolation from other professional groups and subtly taught to guard their own turf and distrust persons from other groups.

7. **More Comprehensive Assessment Information**. In the past, many children were diagnosed, labeled, and placed into special programs on the basis of limited or fragmented information. The law now clearly encourages us to look at the child in a broad context, which means we need data on the "whole child." Instead of conducting only psychological evaluations, which used to consist largely of intelligence tests, we are now also being asked to examine educational, sociocultural, and adaptive behavior characteristics. The gathering of sociocultural and adaptive behavior information is important for examining the child's environment and his or her ability to function in it. This is time consuming and professionals will need to be given time for these assessment activities.

While more comprehensive assessment information is intended to benefit the child, such data—if not gathered, interpreted, and used responsibly—could be extremely detrimental. For example, sociocultural and adaptive behavior information is relatively sensitive. If it is used in "gossip," the child is likely to suffer. Part of the training of all professionals should include the importance of confidentiality.

8. **Wider Acceptance of Different Types of Devices and Procedures**. In the past, data from assessment techniques that were not norm-referenced was suspect. Particularly, information from observation and criterion-referenced devices, as well as from parent interviews, was under fire—even though educators pointed out the limitations of norm-referenced devices, when used alone, in providing the kind of information needed for programming. The new mandates imply that a variety of information sources and assessment techniques is not only allowable, but desirable. When devices complement one another, a combination *can* provide more comprehensive information.

9. **Cultural Pluralism**. For some time, schools have been criticized for recognizing and valuing only one culture. Tests and curricula have been termed "monocultural." Courts have concluded that many of the tests used in the classification and placement process were discriminatory to some children and that they recognized and measured only skills valued by the white majority culture. Many professionals have asserted that the schools are based upon the belief of the superiority of the Anglo- or white majority culture and that the Anglo-conformity theory does not adequately describe the complexities of an American society which contains many ethnic groups. PL 94-142 supports a pluralistic view by ensuring the collection of sociocultural and adaptive behavior information and the examination of norm groups to determine if children like those to be tested were included in the standardization sample. The acceptance of the idea of cultural pluralism requires a change in attitudes of all institutions, agencies and professionals working with young children. Perhaps this legal mandate will accelerate the acceptance process.

10. **The Number of Children Eligible for Services Versus The Number that will Receive Services**. Some children may never receive needed services. More thorough evaluations of each child's skills means assessment will be more time-consuming. Teachers may become frustrated about the time it takes for the child actually to receive help once referred. This frustration may lead to resignation and passivity which, in turn, will decrease the number of referrals. It is also possible that school psychologists and special education directors will recommend that a teacher not refer a child (a) because they don't have appropriate tests or (b) programs are unavailable within the school. Often schools don't want to bear the financial burden of sending a child to another program, and waiting lists are illegal.

Suggested Model for Compliance

In discussing the individual mandates, we have pointed out many of the problems in complying with them. For example, tests do not exist to measure skills in a nondiscriminatory manner for certain types of young handicapped children. It is at this point that many professionals throw

up their hands and say, "I give up; if we don't have the right kinds of tests, then how can we be expected to comply with the law!" Perfection is never easy. Compliance will be difficult, but if we try, it is possible. In this section, a model is presented to help you in complying more responsibly with the mandates. It will help especially with the complexities of the nondiscriminatory testing mandates.

Many times people respond to the mandates by asking some expert to give them a list of nondiscriminatory tests. Given the tests we have and the current level of training of many who select and administer them, this request is impossible to fill. Even when professionals select tests which they consider good, the tests may not be suitable in every situation. For example:

1. Some children may not have had the experiences which are measured on the test.
2. Minority group children were not included in the norm sample.
3. The child speaks Spanish—the test is in English, and you speak only English.
4. The child is physically handicapped, and the test requires motor responses.
5. The examiner has not had proper training.
6. The child is fatigued or angry.
7. The child has limited attending behavior.

As you can see, potential sources of bias are complex and often interact; they include the assessment process, the test itself, the child because of his characteristics, and the examiner because of his skills and attitudes.

Potential sources of discrimination are those things which influence the child's performance and lead to inaccurate results. The Protection in Evaluation section of PL 94-142 was designed to eliminate injustices by ensuring a broad, comprehensive, and accurate picture of each child regardless of race, culture, or handicap. The model which follows helps in obtaining an accurate picture of each child by helping assessment personnel identify the potential sources of bias within the assessment process and then trying to minimize those biases. The model is not new. Many competent clinicians and professionals have been doing what it recommends for years.

The "Model for Identifying and Minimizing Potential Sources of Child Evaluation Bias" was developed jointly by Gloria Harbin, a special educator, and John Brantley, a school psychologist. Dr. Brantley is the past president of the National Assocation of School Psychologists. The joint development of an assessment model by members of two different professional groups is important. If the notion of multi-disciplinary assessment is going to work, models must be designed and accepted by different professional groups. Since its development, this model has been presented to and used by a variety of professionals: school psychologists, special educators, speech and language clinicians, social workers, and physical therapists. This is an important point since everyone on the multi-disciplinary team who assesses the child is responsible for complying with the mandates. These mandates do not just apply to the school psychologist.

The "Model for Identifying and Minimizing Potential Sources of Child Evaluation Bias" contains seven steps between the time a referral is made and a placement is recommended: (1) identify sociocultural information, (2) identify child characteristics, (3) identify test characteristics, (4) select test(s), (5) administer test(s), (6) interpret child's performance, (7) prioritize decisions and recommendations (see Figure 1). Because it is possible for bias to occur during any or all of these steps, we have listed under each step the things one should consider in attempting to identify and then minimize biases. Remember, for some children it will be impossible to eliminate the biases totally. However, this approach facilitates systematic identification and minimization of biases, and it is the most responsible one we know for these children. To make this process easier, to increase communication among multi-disciplinary team members, and to document attempts to meet the intent of the law, checklists have been developed for use in the first six steps of the process. To use these checklists for all children receiving assessment is too time-consuming. However, they should be used with children who pose particular assessment problems and who have the potential of being discriminated against. Frequent examination of the checklists will jog your memory regarding good practice and what should be considered as you test.

How to Use This Model

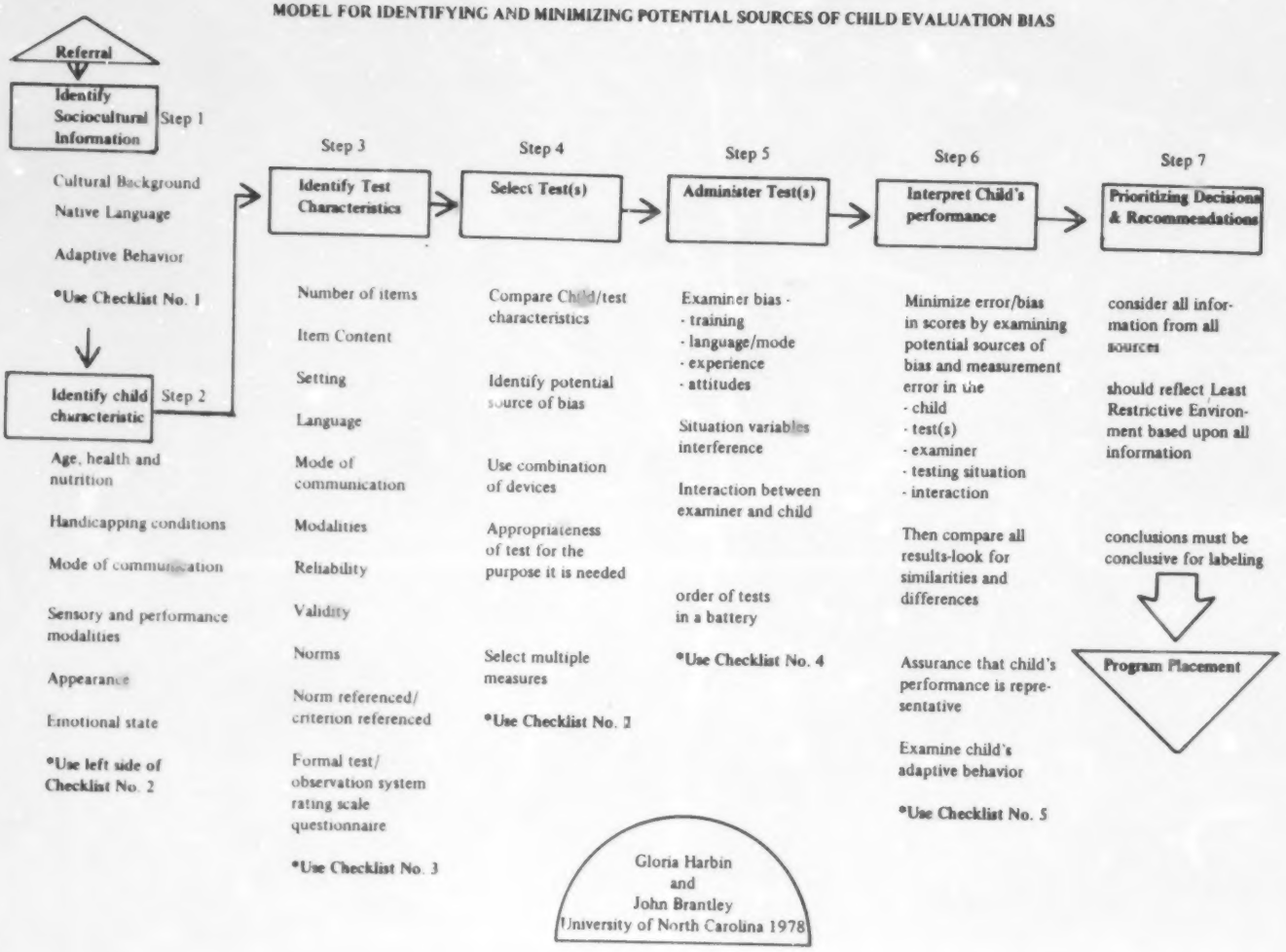
1. Look at the model and the sequence of steps. Notice that Step 1 begins immediately after referral of the child. Also notice that the steps in the model follow the normal assessment process and end with placement of the child.
2. Look at each step. Notice that those things which could be potential sources of bias within that step are listed.
3. To identify and minimize bias for a particular child, use the checklist noted for that step.

Step 1: Identify Sociocultural Information

a. Note the following information for each child: the income, educational, and occupational level of the parents; prior educational and test experience of the child; characteristics of the neighborhood; region of the country; and minority group membership. Could any of these variables influence the test you might select or the child's performance on the test?
b. Identify the child's dominant language.
c. Collect adaptive behavior information (how the child functions in home and community). *Use Checklist No. 1

Figure 1

MODEL FOR IDENTIFYING AND MINIMIZING POTENTIAL SOURCES OF CHILD EVALUATION BIAS



Step 2: Identify Child Characteristics

a. Pay attention to the age, health and nutrition of the child. Do they have any influence on performance?
b. What are the child's handicapping conditions, and how do they affect his performance?
c. What mode(s) of communication does the child use (oral, manual, total, language board)?
d. What sensory and performance modalities are used to receive and understand information as well as express or communicate information?
e. What is the child's physical appearance, and does it influence the way you respond or feel? For example, there have been instances where children were viewed unfavorably because they had a runny nose and were dirty.
f. What is the child's emotional state? If the child is anxious or depressed, for example, it will influence his ability to perform.
*Use left side of Checklist No. 2

Step 3: Identify Test Characteristics

To understand the strengths and the limitations of the tests and other assessment procedures that we use, we must examine the following characteristics:

a. How many items are there? Some tests have too few.
b. What is the item content? Does it measure what you want it to?
c. Is the child's behavior measured in the home, the classroom, or is it generalized to all settings?
d. What language (or languages) is (are) used?
e. In what modes of communication is the test available?
f. What response modalities are used?
g. How adequate is the reliability?
h. How adequate is the validity?
i. What is the size and composition of the norm group? Is it adequate?
j. What type of test is it—norm-referenced, criterion-referenced, observation, Piagetian?
*Use Checklist No 3

Step 4: Select Test(s)

a. PL 94-142 encourages individualization in the selection of tests and assessment techniques. Therefore, have you matched the characteristics of the information needed on the child with the characteristics of and information provided by the test?
b. Have you identified potential sources of bias with a particular test and decided how to minimize them?
c. Have you used a combination of devices? The use of criterion-referenced devices and observation can supplement data from norm-referenced tests.
d. Have you made sure the test is appropriate for the purpose for which you are using it? If your purpose is placement, you must have data from some norm-referenced device.
e. Have you selected more than one measure? Make sure measures are educationally relevant and are measuring all areas of suspected disability.
*Use Checklist No. 2

Step 5: Administer Test(s)

Have you checked for potential sources of bias in the following areas:

a. The examiner?
b. The testing situation or environment?
c. The interaction between the child and the examiner?
d. The order of subscales in a test and/or the order of tests in a battery? For example, a shy preschool child may not perform gross motor or expressive language tasks as well in the beginning of a testing situation as he would once he feels more comfortable.
*Use Checklist No. 4

Step 6: Interpret Child's Performance

a. Have you examined, considered, and synthesized all of the information obtained and uncovered potential sources of bias which could result in measurement error?
b. Have you compared results of multiple measures and sources, and looked for similarities and differences?
c. How do these results compare to the measure of the child's adaptive behavior? A child must be delayed in both intellectual and adaptive behavior before he can be classified as mentally retarded.
d. What indications do you have that these results are representative of the child's behavior? For example, did you do more than one classroom observation?
*Use Checklist No 5

Step 7: Prioritizing Decisions and Recommendations

a. Have you considered all data from all sources?
b. Have you selected the least restrictive environment? The most appropriate least restrictive environment is not always the regular classroom.
c. Is your decision conclusive enough for labeling?

Checklist 1

Form for Documenting Potential Sources and Effects of Sociocultural and Language Bias

Potential Source of Bias **Potential Influence on Evaluation**

Sociocultural

____ Income level of parents
____ Educational level of parents
____ Occupational level of parents
____ Prior educational experience of child
____ Prior test experience of child
____ Regional or neighborhood characteristics
____ Minority group membership characteristics

_____
_____
_____

Other Considerations

_____
_____
_____

Dominant Language (if other than English)

Name of language measure used to determine language dominance

_____

Other source (specify)

_____

Language or dialect indicated

_____

Name of examiner or translator fluent in language or dialect indicated

_____

Adaptive Behavior

Name of adaptive behavior measure

_____

Results of adaptive behavior assessment

_____

Does this device measure child's ability in the home and community?

_____

Gloria Harbin
and
John Brantley
University of North Carolina 1978

Checklist 2

MATCH SHEET FOR TEST SELECTION

CHILD	TEST	DEGREE OF BIAS: NONE, SOME, CONSIDERABLE
Age, Health, Nutrition	Age	
Handicapping Condition	Appropriate With What Handicapping Conditions	
Communication Modalities Child-Test (Modalities Used) Child-Examiner (Mode of Communication)	Communication Modalities Child-Test (Modalities Used) Child-Examiner (Mode Communication)	
Cultural Background	Norm/Reference Group	
Native Language	Language Availabilities	
Educational Exposure	Assumed Experiences	
Information needed about performance in which setting (e.g., home, classroom)	Setting	
Areas to be tested	Purpose/Content areas	

* * * * * * * * * * * * * * * * * * *

Type of Bias	Procedure to Minimize Bias

Gloria Harbin
and
John Brantley
University of North Carolina 1978

Checklist 3

Test Characteristic Work Sheet

Name of Measure _____

Purpose _____
Language Alternatives _____
Communication Modalities
Child-Test _____
Child-Examiner _____
Norms:
Age Range _____
Sex _____
Grade _____
Race/Ethnic Status _____
Socioeconomic Status _____
Geographic Region _____
Urban/Rural _____
Local Norms _____
Setting _____
Examiner Qualifications _____
Test Characteristics
Reliability _____ _____ Internal
_____ _____ T-RT
Validity _____ Face
_____ Content
_____ Concurrent
_____ Predictive
_____ Construct
Item Relevance _____
Type of Measure _____
Scoring Criteria _____
_____
Types of Scores _____
_____

Transfer relevant information to Checklist No. 2; match sheet for test selections

Gloria Harbin
and
John Brantley
University of North Carolina 1978

Checklist 4

Checklist for Identification
of Potential Bias During Administration

Potential Examiner Bias

_____ Training (philosophy, skills for all ages and all handicapping conditions)
_____ Language/Mode (what language and mode of communication are needed by the examiner with this child)
_____ Experiences testing similar types of children: ages, cultural groups, handicapping conditions
_____ Attitudes toward particular cultural groups and about nonbiased assessment
_____ Knowledge of alternative measures

Situational Interference

_____ Time of day
_____ Distractions
_____ Testing materials (color, size, etc.)
_____ Inadvertent use of cues such as position cues/position of materials
_____ Length of session
_____ Comfort and accessibility of materials
_____ Order of assessment activities

Interaction between Examiner and Child/Respondent

_____ Establishment of rapport
_____ Obtaining and maintaining attending behavior
_____ Initial success or failure of child
_____ Maintain child's optimum effort
_____ Adequate communication (mode, manner, language)
_____ Adequate modeling or demonstration
_____ Dress and/or mannerisms of examiner (distracting)

Order of Test

_____ Did the order of tests given or items given influence the child's performance

xxxx

Gloria Harbin
and
John Brantley
University of North Carolina 1978

Gloria Harbin
and
John Brantley
University of North Carolina 1978

Checklist 5

Checklist for Minimizing Bias During Interpretation of Results

Examine Child's Scores

_____ Compare them to the adaptive behavior information
_____ Look for characteristics of the child which might bias or influence the results such as:
- _____ native language
- _____ age, health, nutrition
- _____ handicapping conditions
- _____ mode of communication
- _____ sensory and performance modalities
- _____ emotional state

_____ Look for characteristics of the tests and techniques which might bias or influence the results, such as:
- _____ purpose
- _____ communication modalities (a) child-test (b) child-examiner
- _____ norms
- _____ reliability and validity
- _____ type of measure
- _____ relevance of items
- _____ scoring criteria
- _____ type of scores

_____ Look for characteristics of the examiners which might bias or influence the results, such as:
- _____ appropriate training
- _____ communication mode and language
- _____ previous experience
- _____ attitudes
- _____ skills
- _____ knowledge

_____ Look for conditions within the assessment situations which might bias the performance, such as:
- _____ time of day
- _____ distractions
- _____ testing materials
- _____ inappropriate use of cues
- _____ length of session
- _____ comfort and accessibility of materials
- _____ order of assessment activities

_____ Look for conditions between the examiners and child which might bias the performance
- _____ rapport
- _____ attending behavior
- _____ initial success or failure
- _____ maintaining responding behavior
- _____ communication
- _____ directions, modeling or demonstrating
- _____ dress and/or mannerisms
- _____ child's appearance

_____ Try to determine if the child's performance is representative and/or approximates his/her potential

_____ Compare the results of multiple measures

References

American Psychological Association, American Educational Research Association and National Council on Measurement in Education. *Standards for Educational and Psychological Tests.* Washington, D.C.: American Psychological Association, Inc., 1974.

Brantley, John C. "Non-Discriminatory Evaluation." In A. Turnbull, B. Strickland, and J. Brantley (Eds.), *Developing and Implementing IEPs.* Columbus: Charles E. Merrill, 1978.

Department of Health, Education, and Welfare: Office of Education. "Education of Handicapped Children: Implementation of Part B of the Education of the Handicapped Act." *Federal Register* (Pt. 2) August 23, 1977, pp. 42473-42518.

The Education for All Handicapped Children Act, Public Law No. 94-142, 20 U.S.C. 1402, *et seq.*

Hoeussermann, Else. *Developmental Potential of Preschool Children.* New York: Grune and Stratton, 1958.

Journal of School Psychology. "A Special Issue: School Psychology and Early Childhood Education." 15 (1977).

Mercer, J. R. *Labeling the Mentally Retarded: Clinical and Social System Perspectives on Mental Retardation.* Berkeley, California: University of California Press, 1973.

Morrow, H. W., and Coulter, W. A. (Eds.). *The Concept and Measurement of Adaptive Behavior Within the Scope of Psychological Assessment.* New York: Grune and Stratton Inc., 1979.

Oakland, Thomas (Ed.). *The Psychological and Educational Assessment of Minority Children.* New York: Brunner/Mazel, Inc., 1977.

Office of Civil Rights. *Memorandum to State and Local Education Agencies on Elimination of Discrimination with the Assignment of Children to Special Education Classes for the Mentally Retarded.* 1972.

Salvia, John, and Ysseldyke, James. *Assessment in Special Education.* Boston: Houghton Mifflin Company, 1978.

Sattler, J. M. "Intelligence Testing of Ethnic Minority-Group and Culturally Disadvantaged Children." In L. Mann and D. Sabatino (Eds.), *The First Review of Special Education* (Vol. 2). Philadelphia: JSE Press, 1973, 161-201.

Silverman, R., Noa J., and Russell, R. *Oral Language Test for Bilingual Students: An Evaluation of Language Dominance and Proficiency Instruments.* Portland: Northwest Regional Educational Laboratory, 1976.

CHAPTER 10

Adaptations in Measurement Procedures: Should you make alterations for handicapped children?

by Rebecca Fewell DuBose

The 1970's will be viewed by educational historians as an important decade in education. Important changes are occurring because the doors of public schools have been opened to all children regardless of the extent of their handicapping conditions. Their presence has required many changes in the delivery of educational services.

Personnel needs are quite different from the needs previously described in special education. Job descriptions require teachers certified in multiple disabilities, physical therapists, language specialists, occupational and recreational therapist, and nurses.

School buildings constructed in the 1970's are built with handicapped children in mind. The Harris Hillman School, located in Nashville, was opened in 1975 as an entire school specifically designed for multiply handicapped children; it has very little in common with traditionally constructed school buildings of a decade earlier.

The curriculum for educating the school children of the 1970's has likewise experienced dramatic changes. A friend of mine, as she supervised the first day of school for a group of severely handicapped children, painted a very real picture of the new needs:

> *When the bus arrived she saw no happy clean faces of neatly dressed children eagerly anticipating their first day in school. Rather she saw only silent children waiting to be carried off the bus. It became apparent immediately that toilet skills were not in their repertoires. The school and its staff were not prepared. The supervisor hastily called the superintendent and exclaimed, "Forget that order of pencils, paper, scissors, and paste. Within the next hour send me Pampers–big ones, plastic bags, washclothes, floor mats and Hogg chairs." The red tape of purchase orders in that particular school system has never been cut so quickly as it was on that day: the impact of providing educational services to severely handicapped children had been felt (DuBose, 1978, p. 14).*

Another important thrust in educational services in the 1970's is measuring progress. We must show that we are affecting the lives of children in a positive way. We can no longer simply use adjectives to describe a child's performance. Programs can no longer expect federal assistance based on past laurels, or on the basis of cursory descriptions of the good things that are happening to the children in a program. We must quantify our reports and precisely report how much progress has occurred. For this reason, we must be very careful in the selection and administration of measurements.

I have been asked to address the issues surrounding adaptations in testing. In doing so, I would like to briefly review our current models for assessment, note the problems confronting today's examiners, and suggest adaptations.

Historical Overview of Test Development. Adaptations in testing can be traced to the genesis of the testing movement. One of the early leaders in the field of assessment, Sir Francis Galton, was convinced that the keener the senses, the larger the field upon which judgment and intellect can act. To test this, Galton developed a series of tests of sensory discrimination which he used to define the intellectual abilities of individuals. Alfred Binet felt these and other available tests were too sensory and focused on unduly simple abilities. Together with Simon he published a scale that had a greater proportion of verbal content in contrast to the sensory perceptual tasks included in other tests. The basic content of almost all mental measures we use today are adaptations of the original Binet scales.

Present Models for Assessment. The traditional measures most of you are familiar with and are using were designed for use with normal children. Many were standardized on youngsters without known impairments. Based on the norms provided, the examiner can compare the performance of the child being tested to that of the appropriate normative population.

Another form of measurement, criterion-referenced testing, is rapidly gaining status. An individual's performance is evaluated in terms of an absolute or specific level of mastery, or criterion, that has been previously set. Criterion-referenced tests are often directly linked to specific instructional objectives and for this reason, are particularly relevant to the kind of programming that must be planned for today's children. Unfortunately, even criterion-referenced tests must also be adapted for many severely involved children.

Today's needs call for both kinds of information:

we must *identify* and, in some cases, *classify* children in order to serve them. And we must identify what they know, what they need to know, and how they can best be taught in order to serve them. To meet these needs, a model of assessment that permits the examiner to use both norm-referenced and criterion-referenced testing is helpful. Using both kinds of measurement models, the examiner can carefully study the differences within the child while at the same time having a way for determining the child's performance relative to children of the same age group. Within this framework, an examiner can adapt tests and testing conditions in an effort to test the child's limits and to identify the conditions under which he can or will respond. These samples of behavior provide far more functional information than can be gathered through traditional testing methodology alone.

A word of caution is needed. Two things must be very clear. Once the examiner decides to "adapt" a test by changing any part or procedure, the standardization has been violated and the child's performance cannot be compared to the normative sample with any degree of assurance that the findings are accurate. If an examiner changes the administration of an item, it is imperative that he state why it was done and exactly what was done. This will enable the next examiner to test the child under similar conditions.

Before we begin to make changes in the tests currently available, it is important that we carefully consider why some tests or test items are not appropriate for the children we are serving today.

Problems in Available Tests. Most of the test available today were not developed with severely impaired persons in mind, and in many cases, the instruments are not appropriate or valid measures (DuBose, 1978). Many instruments that are used: (1) lack stimulus value, a quality that is important in motivating the handicapped child; (2) contain items that are time-bound and are not recognized by today's children; (3) require language skills that are likely to be depressed in handicapped children; (4) require fine motor manipulations that are frequently impaired in handicapped children; and (5) require speed in performance, a concept difficult to convey to handicapped children. Perhaps the most significant shortcoming in the tests available is the limited amount of data they supply which can be translated into appropriate goals, objectives, and teaching strategies.

Problems in the Assessment Environment. We have made strides in correcting some of the problems inherent in tests, but we have done very little to change the environment in which we test. For the most part, testing continues to take place in barren rooms, housing only a table, two chairs, a test kit, and a stopwatch. Ironically, this sterile environment is far removed from the environments in which the behaviors we are to assess naturally occur.

The context in which behavior takes place has an important effect on performance. Bortner and Birch (1970) noted the glaring differences that occur in estimates of potential when alterations are made in performance conditions. Individuals do not usually act independently of outside forces in any given situation, but continually respond to situational factors that may or may not be present to the casual observer (Wallace and Larsen, 1978).

It seems imperative that we give very serious study to the assessment milieu; that we expand and adapt our testing typologies with the same zest that we have invested in the study of measurement devices.

We have focused on the problems in testing and in testing environments. It is now appropriate that we focus on possible directions that may prove fruitful in our search for more meaningful assessments.

Adaptations in Testing Procedures

Adaptations in testing permit the examiner to be flexible in his approach, and sensitive to the child's needs. The examiner may need to establish rapport and a communication system that involves physical contact, particularly in the case of very young or severely involved children. Several years ago I had the pleasure of working on an assessment team with a noted psychologist, Dr. Norman Buktenica. He is an expert in establishing rapport with handicapped children. I shall never forget one very moving experience. Before he began administering test items to a totally deaf and blind sixteen year old who had spent his last ten years in an institution for the retarded, Dr. Buktenica communicated to the boy that he wanted to shake his hand after he had completed each task. This was done by giving the boy a simple task to do, then immediately reaching for the boy's hand once it was obvious the task was complete. The system was quickly established and was used with joy by both throughout the assessment period. The addition of this simple handshake would be considered an "adaptation" and a violation of norm procedures. Without this simple communicative exchange and the adaptations of the tasks presented to the teenager, our team would never have gathered the necessary data that freed this youngster from institutional bondage and paved the way for his entrance into an education system and later into successful employment in his home community.

Most tests provide one or two sample items that can be used in demonstrating to the child how he is to respond. Seldom are these sample items adequate to ensure that the handicapped child understands the expectations. At times it is necessary to design a simple task that requires the child to respond in the same manner as the test will require. Once the response paradigm has been learned, the examiner can proceed to the sample items with a greater likelihood of the child understanding the response pattern. This pretesting activity would also violate the administration of some tests and must be noted as an adaptation.

Frequently, it is necessary to alter the response pattern more dramatically. The correct response on a particular test might be to respond by verbalizing a correct discrimination. A child may not be able to verbalize the responses, but, when allowed to point to the objects, he can demonstrate his awareness of the differences between them. Such changes clearly alter the task and must

be reported. Yet by making the alteration, the examiner can discover the level on which the child can respond most effectively and can provide more sensitive data as to the child's instructional needs and learning level. Carrow's (1972) hierarchy of response levels are noted in Table 1.

Adaptations using a test-teach-test model permit the gathering of information that is relevant to the child's individual educational plan. If the child fails a task, yet is shown how to perform the task, one can observe his imitation skills or perhaps, at a later time, his memory. Problem-solving skills, approach to tasks, use of resources, attention span, response to varying materials and media, and generalization skills are among the things an alert, skilled examiner can note in a test-teach-test mode.

Test materials are notoriously boring. Most do not invite interaction and consequently evoke minimal responses. By changing the material and increasing the stimulus quality of the task, one can determine whether the child understands the task, but failed to respond because of the material. Additionally, one begins to identify the qualities in materials that evoke preferred responses from each child. The point is not to change the concept that is being examined, but to increase the stimulus value of the materials being used.

Adaptations in the Testing Milieu

By adapting the environment, one can maximize the

TABLE 1

Relationship Between Piaget's Stages of Sensorimotor Development And Bates' Model of Communicative Development*

Piaget's (1952) Stages of Sensorimotor Development	Bates' (1976) Proposed Sequence of Communication	Exemplifying Behavior	Child's Communicative Intention	Adult's Interpretation
Use of Reflexes	Perlocutionary Acts	Crying	None	He's hungry, he wants to be picked up, etc.
Primary Circular Reactions	Perlocutionary Acts	Repetitive sucking movements	None	He's hungry, he wants to have his bottle, etc.
Secondary Circular Reactions	Perlocutionary Acts	Vigorous arm movements, smiling, bodily agitation, etc. upon seeing an adult perform an interesting behavior	"Do it again"	He wants me to continue making the funny face.
Coordination of Secondary Circular Reactions	Illocutionary Acts	Extending arms to be picked up	"Pick me up"	He wants me to pick him up.
Tertiary Circular Reactions	Illocutionary Acts	Handing a wind-up toy to an adult	"Make it work"	He wants me to wind it up so he can see it go.
Invention of New Means Through Mental Combinations	Locutionary Acts	Child says "More juice"	"I want some more juice"	He wants something else to drink.

*Dunst, 1978

likelihood of appropriate responses. A simple example will illustrate the point. The assessment of language most frequently takes place in a traditional testing room with the child responding to a given set of questions or stimulus events. This might yield some correct information about the child's language. In my experience with the assessment of language in handicapped children, however, I have found out far more by placing the child in a stimulating environment where he can be an actor in an arena. In this situation I can observe whether he understands instructions, plays functionally, follows commands, and verbalizes about events. This is even more true for teaching language. We have been far more successful teaching language to some children in a gym filled with swings, ladders, trampolines and jungle gyms, than in sterile, isolated language areas occupied by a table, two chairs, picture cards, and M &M's.

Within the milieu is the examiner. It is quite true that in the real world we must relate to many persons. We interact differently with each person. To capture the responses of the child, it seems appropriate to view his interaction with several persons, including family members. In recent years, I have been genuinely impressed by the exchanges between handicapped children and their siblings. We now include siblings as a very real part of our assessment of handicapped children. From siblings we have learned new ways to approach the child, ways to reinforce him, ways to test him and to teach him.

Testing must be adapted to fit into the real world of the child. The materials we use, the procedures we follow, the human exchanges that transpire must be considered in light of our ultimate goal: that of facilitating the nurturing of children.

References

Bortner, M., and Birch, H. C. "Cognitive Capacity and Cognitive Competency." *American Journal of Mental Deficiency* 75 (1970): 735-744.

Carrow, E. "Assessment of Speech and Language in Children." In *Language Intervention With the Retarded: Developmental Strategies*, edited by J. E. McLean, D. E. Yoder, and R. L. Schiefelbusch. Baltimore: University Park Press, 1972.

DuBose, R. "Identification." In *Systematic Instruction of the Moderately and Severely Handicapped*, edited by M. E. Snell. Columbia: Charles E. Merrill, 1978.

Wallace, G., and Larsen, S. C. *Educational Assessment of Learning Problems: Testing for Teaching*. Boston: Allyn and Bacon, 1978.

CHAPTER 11

Documenting Child Progress: How do you choose data collection designs and instruments?

by Robert Sheehan

When we refer to the relationship between data collection designs and instrument selection, we are really referring to the broader issue of documenting children's progress. The present discussion is designed to suggest decision making strategies which can be followed by early childhood special educators in the documentation of children's progress. At the very outset, I must say that you cannot select a data collection design without considering your instruments, nor should you select instruments without some foresight and anticipation of your design.

Allow me to clarify what I mean by the term "instrument." I prefer not to define "instrument" in a restrictive way (for example, a tool published by a testing corporation). Instruments are any attempts to gather information systematically. An instrument is a data collection guide which: (a) measures what it purports to measure (validity), and (b) insures that information yielded by one person using it is qualitatively similar to the information produced by another person using the same instrument on the same subject (reliability). This working definition suggests that the term "instrument" can be used for rating scales, observation and interview schedules, questionnaires, tests which measure a single dimension, tests which assess a variety of developmental or other areas, and much more.

Attention is usually directed toward the more formal tests. This is not because other data collection procedures are less important. Rather, it seems that when child progress data are gathered for evaluation purposes, we in early childhood special education turn toward data collection procedures which have the broadest applicability and are the most widely understood by parents, teachers, administrative personnel, and funding agencies. These procedures are most often the formal tests, but there are many instances when other types of instrumentation are excellent techniques for demonstrating the progress of children.

Underlying this discussion of types of instruments is my first point regarding the relationship between data collection, design and instrument selection. An attempt should be made to gather data which present an argument useful for decision making about programs for children. The argument which you intend to make should be one factor which affects your selection of both the data collection design and the instruments. For example, if you are working with children with communication problems and you wish to argue that you have had an impact on their communication skills, then select a data collection design and instruments which provide data for or against this argument. Other areas of development should be considered if they are within your goals, but do not in any way neglect the gathering of data about children's communication development.

My colleague, Dr. Barbara Keogh, and I (Sheehan and Keogh)[1] completed a review of the current (state-of-the-art) way child progress is documented in early childhood special education as part of our involvement with Project REACH, the BEH funded Early Childhood Research Institute at UCLA. One message which is apparent from the development of this review is that there is variety in the type of arguments which programs present in documenting child progress. The variety occurs because of the tremendous range of handicapping conditions with which programs are dealing, and because of the differences in philosophies and goals among projects. Consequently, there are several different data collection designs employed to support these arguments.

Data Collection Designs

Our review has indicated that there are five categories of quantitative data collection designs currently in use in early childhood special education programs.[2] Several of these are similar to designs discussed by Campbell and Stanley (1966). Other designs are either new or significant variations of the traditional designs. The first design is the *one group pretest posttest design.* This is a design in which two assessments of handicapped children are made to gather data to present as documentation of

---

[1] R. Sheehan and B. K. Keogh. "Documentation of Progress" (REACH Working Paper). Los Angeles: University of California, Graduate School of Education, Fall, 1978.

[2] Attention is focused in this chapter on data collection designs which yield quantitative evidence of the progress of children using scores on instruments as the evidence. Other more qualitative techniques such as case studies are not discussed. In fact, our review efforts have failed to identify any qualitative designs in use in early childhood special education.

child progress. These assessments most logically are made prior to and following intervention. The critical features of this design are that these assessments are only taken of handicapped children receiving the educational intervention and that only two assessments are taken.

The second design is a *time series design* in which a number of assessments are made on handicapped children and then presented as evidence of child progress. The critical features of this design are that the assessments are made frequently (for example, daily, weekly, monthly) and that the assessments are made only on the children receiving the intervention.

The third design, which is used quite frequently in general educational research but much more infrequently in early childhood special education programs, is the *non-equivalent control group design*. It consists of two assessments being taken on two groups of children. The assessments are usually taken prior to and following intervention and the two groups of children include the children receiving the intervention and another group of children who are judged similar to those receiving the intervention. This design is not used frequently due to the difficulty in obtaining comparison children who are similar but are not also receiving services.

The fourth and fifth types of designs are being used in some form or another by a number of early childhood special education programs. Our work at UCLA and the work of Huron Associates of Cambridge is designed, in part, to test the validity of designs like these - which are called *prediction designs*. They are similar to designs used in evaluating Title I Programs (Tallmadge, 1977). The fourth type of design is the *one-shot prediction design*. It has two critical features. The first is that the design involves only one assessment of handicapped children's progress. The second is that data are collected only on the children receiving the educational intervention and these data are then compared to a hypothetical comparison group which is the norming population of the instrument. In essence, the comparison is made in the prediction design to an estimate of what children's progress *would have been in the absence of intervention*.

The fifth and final category of designs is what we call the *pretest posttest prediction design*. This is different from the one-shot prediction design in that two assessments of handicapped children's progress are made rather than one. These assessments occur usually prior to and following intervention. As in the previous design, data are collected only on children receiving the intervention. Once again, handicapped children's performance is compared to an estimate of what the performance would be in the absence of intervention. This estimate is derived by computing a predicted posttest score based upon children's pretest scores.

I would like to mention one postscript to these five categories of designs currently in use in early childhood special education programs. We did not find any programs which utilized random assignment of subjects to treatment and control groups. With the implementation of P. L. 94-142, various state mandates, and the ethical concerns of professionals in the field regarding deprivation of services, random assignment to treatment and no-treatment groups is presently not an option. Our efforts, therefore, have turned to a search for alternatives to the randomly assigned control group design.

Two Conceptual Systems

During our work at UCLA, we found ourselves asking: "What are the critical features of these designs?" We identified two dimensions which early childhood special educators must consider in selecting appropriate designs.

The first dimension concerns *comparisons* made in the designs. That is, to what is child progress being compared? We identified four possible answers to this question: (1) self comparison; (2) a randomly assigned control group; (3) a quasi-control group: (4) a statistical or predicted control group.

The second dimension concerns *frequency of assessment*. That is, how often are assessments made? We found that most projects either conducted: (1) a single assessment; or (2) two assessment; or (3) multiple assessments.

These two dimensions when considered together yield the matrix shown in Table 1. Of the twelve cells in the matrix, only five were found to be in use in our review of current practices in early childhood special education projects.

TABLE 1

Data Collection Designs in Early Childhood Special Education

		Frequency of Assessment		
		Single	Two Times	Multiple
Type of Comparison	Self-Comparison		1	2
	Random Control			
	Quasi-Control		3	
	Prediction	4	5	

1. The One-Group Pretest-Posttest Design
2. Time Series Design
3. The Non-Equivalent Control Group Design
4. The One-Shot Prediction Design
5. The Pretest-Posttest Prediction Design

A third type of question, or dimension, is of equal value in making decisions about documenting child progress: "Considering the children with whom you are working and your program's philosophy and goals, what type of argument do you want to make regarding children's growth and development?" This question, in addition to those

about *type of comparison* and *frequency of assessment* (Table 1), should ultimately guide your selection of both instruments and data collection designs.

I have only begun to consider (Sheehan, 1979)[3] possible answers to this third question. This discussion therefore, should be considered tentative. It seems to me that with children who are developmentally delayed, the answer might be, "We want to argue for normal development across all developmental areas." With children who exhibit an uneven developmental profile, such as learning disabled children, it may be: "We want to argue for normal development in the affected areas as a result of our program's efforts." On the other hand, with children whose handicapping conditions are quite involved (for example, profoundly handicapped children), the answer may be: "We want to demonstrate any progress at all in the beginning and then develop and constantly update our expectations based on prior performance."

A fourth answer to the question might come from programs serving children with a specific handicap in one or several areas and for which normal development is unrealistic. For example, deaf children's communication skills might always differ from the hearing child, and we might want to argue for limited growth in these areas and normal growth and development in other areas.

The addition of this dimension leads to a new matrix: Table 2. Bear in mind the tentative nature of this matrix, and use it primarily as a guide to help you understand your own decision making process.

[3] R. Sheehan. "Documenting Children's Progress." Council for Exceptional Children. Dallas, Texas, April, 1979.

Integrating Data Collection Designs and Instruments

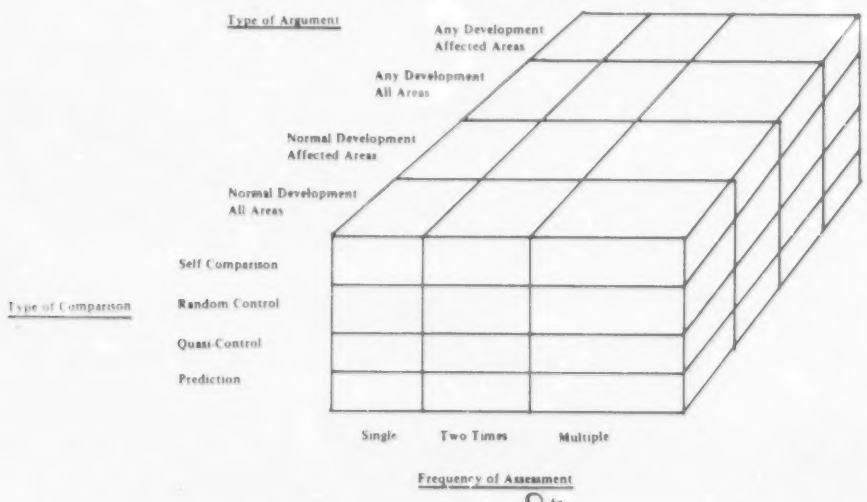
The problem now becomes, "How do we translate a matrix such as this into the search for appropriate instruments." Let us review the three questions which make up the decision making process.

1. Considering the children with whom you are working, and your program's philosophy and goals, what type of argument do you want to make with your documentation efforts regarding children's growth and development?
2. What type of comparison must you make in your design to insure that your results did not occur for any other reason than your early intervention?
3. How frequently must you gather data on children's progress?

It seems to me that answering these questions in this order, puts you on the right track to search for instruments. There is no single answer about what instrument to choose or what design to select. The process, though, by which you as decision makers attempt to answer the questions is of critical importance. As decision makers, you must juggle the complex set of variables to arrive eventually at the best documentation method for your purposes. I must be very frank and state that some of the instruments currently in the field and some of the data collection designs currently in use raise more questions than answers. Their adequacy for documentation purposes has not been thoroughly investigated. Be aware of possible shortcomings as you make your decisions, but do not give up on the struggle to find the best documentation strategy. As the field evolves, our views of it and our knowledge base are becoming systematically more refined.

TABLE 2

Decision Making in Documenting Child Progess in Early Childhood Special Education



87

References

Campbell, D. T., and Stanley, J. C. *Experimental and Quasi-Experimental Designs for Research.* Chicago, Ill.: Rand McNally, 1966.

Sheehan, R. "Documenting Children's Progress." Council For Exceptional Children. Dallas, Texas, April, 1979.

Sheehan, R., and Keogh, B. K. "Documentation of Progress." (REACH Working Paper) Los Angeles: University of California, Graduate School of Education, Fall, 1978.

Tallmadge, K. G. *The Joint Dissemination Review Panel IDEABOOK.* Washington, D. C.: U. S. Department of Health, Education and Welfare, National Institute of Education, U. S. Office of Education, 1977. (Contract No. NIE-IA-7706.)

ABOUT THE EDITOR

Talbot Black received his BA from Phillips University in Enid, Oklahoma, and his MA from George Peabody College for Teachers in Nashville, Tennessee. He has worked at the Technical Assistance Development System (TADS) since it began in 1971. He is presently Associate Director for TADS. His professional interests include: early education programs for the handicapped and technical assistance and support systems.